Institut d'études politiques de Paris
ÉCOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'économie
Doctorat en sciences économiques

In Search of Frictions

Clément MAZET-SONILHAC

Thesis supervised by

Thomas CHANEY, Professeur des Universités, Sciences Po

defended on June 25, 2021

Jury

Mrs. M. Cecilia BUSTAMENTE, Professor of Economics, University of Maryland *(reviewer)*
Mr. Thomas CHANEY, Professeur des Universités, Sciences Po
Mr. Thierry MAYER, Professeur des Universités, Sciences Po
Mrs. Isabelle MÉJEAN, Professor of Economics, Ecole Polytechnique *(reviewer)*

à Alberte, Suzanne, André et Fernand.

Acknowledgements

I owe an enormous debt of gratitude to my advisor, Thomas Chaney, for his trust, availability, and tremendous support. Thomas is the most passionate and enthusiastic researcher I have ever met. His insatiable curiosity and creative approach to research are extremely contagious. They have inspired me deeply during all these years and will continue to do so. Thomas always encouraged me to explore my own research ideas, no matter how dubious they seemed at first glance, while guiding me with his invaluable insights. I am proud – and fortunate – to have worked under his supervision.

I am also extremely grateful to Thierry Mayer for his support throughout my thesis. Having the chance to collaborate with him was an immense honor and – I must say – an intimidating experience. He guided me through the vast literature in international trade and the estimation of gravity equations, which was decisive for my work and my development as a researcher. Thierry also taught me that bouldering at Fontainebleau is an effective way to escape the tumultuous world of research for a weekend, which was often needed.

I want to thank Maria Cecilia Bustamante and Isabelle Méjean for accepting to read and review my dissertation. It is a great honor for me and of tremendous benefit to my research to have them in my thesis committee.

I am indebted to my co-authors Yasser Boualam, Anne Duquerroy, Jean-Stéphane Mésonnier and Daniel Paravisini. This adventure would not have been as rich and fruitful without them. Special thanks to Clément Malgouyres, who passed on to me his pure and visceral love for economics. I have never met anyone so good – and so obsessed – with micro data, and I am proud of what we have achieved together. I would love to continue to be woken up in the middle of the night by his flurry of messages about CES production functions.

I discovered economics in the thrilling intellectual atmosphere of the economics department at Sciences Po and I am grateful to the entire faculty there. I would especially like to thank Florian Oswald and Michele Fioretti. Florian introduced me to the joy of `Julia` and gave

me the opportunity to teach his Numerical Methods class. Intense discussions with Michele gave me both ideas and confidence; I am proud and excited to share research projects with him. I also feel lucky to have had the opportunity to interact with many brilliant PhD students at Sciences Po and I would like to thank them for the inspiring exchanges we had along the way. Here are some, though I am afraid not all, of the people that were part of these interactions: Tyler Abbot, Vladimir Avetian, Olivier Cassagneau-Francis, Sophie Cetre, Pierre Cotterlaz-Carraz, Cyril Couaillier, Nicolo Dalvit, Samuel Delpeuch, Pierre Deschamps, Edgard Dewitte, Jean-Benoît Eyméoud, Etienne Fize, Arthur Guillouzouic-le-Corff, Dorian Henricot, Jean-Louis Keene, Gustave Kenedi, Alaïs Martin-Baillon, Julia Mink, Elisa Mougin, Ludovic Panon, Asem Patel, Stefan Pauly, Jan Sonntag, Charles Louis-Sidois, Camille Urvoy, Paul Vertier and Max Viskanic. I would also like to warmly thank Pilar Calvo, Hadjila Nezlioui Serraz, Claudine Lamaze and Sandrine LeGoff for their invaluable help at all stages of my thesis.

A significant part of this journey took place at Banque de France, at the Service des Analyses Microéconomiques that welcomed me five years ago. I am deeply thankful to Frédérique Savignac, Erwan Gautier and Vincent Bignon for their constant support throughout these years. This thesis would not have been possible without them. I am grateful to Nicoletta Berardi for having kindly guided my first steps in this crazy research world. I am also grateful to Mathias Lé for our impassioned discussions and debriefings of the Lakers games and for his gentle taunts when I experienced dreadful doubts about my thesis. I would like to thank Laurent Baudry for his tremendous help with the data and his attention to detail that inspired me so much. I am also grateful to Pamfili Antipa, Clémence Berson, Antoine Berthou, Jocelyn Boussard, Simon Bunel, Juan Carluccio, Barbara Castillo Rico, Audrey Deffaux, Françoise Drumetz, Bertrand Garbinti, Paul Hubert, Lisa Kerdelhué, Pauline Lesterquy, Noëmie Lisack, Jérémi Montornès, Sandra Nevoux, Lionel Potier, Loriane Py, Aurélie Sotura, Gabriel Smagghue, Sylvie Tarrieu and Riccardo Zago for the endless discussions that greatly contributed to this thesis. I would also like to thank everyone at Collège de France. I am particularly honored to have had the opportunity to meet and work with Philippe Aghion, who welcomed me warmly into this prestigious institution.

I am also grateful to Frédéric Lordon. Had I never met him, I probably would not have started a thesis. The reading of his books, and his kind advice made me completely change the course of my life almost ten years ago. Frédéric awoke in me an intellectual enthusiasm for social sciences that never left. I hope he will forgive me for acknowledging him in a thesis that begins with a quote from George J. Stigler. As a *forker*, I owe him a lot.

I could not have made it through these years of academic joy and despair without a bunch

of incredible friends. Malo EG and My-Lan (and now Thao!) who kept reminding me that simple things – randomly: the Olympique de Marseille, a fresh beer or Keny Arkana – were more important than all this mess. Stéphane and Verity who made me travel from Inverness to Lourmarin. Nazim for the epic ball games and those marvelous, existential and fiery debates at *194 DMS*. Édouard and Antonin, who brightened my everyday life with their caustic humor and profound intelligence. And thanks to the infamous *Crook Chess Club* for the dreams haunted by pawns and knights. Because some places are as comforting as good friends: I am grateful to Ceillac, where I could escape to the summits and Ladon. A substantial part of my thesis was written in front of those magnificent birch trees.

I would like to thank my beloved family and especially my parents, Cathy and Yves, for their unconditional love and support. They gave me the strength and confidence to get here.

Finally, all this would not make any sense without Lydie. She was by my side through the *very* ups and the *very* downs, when her sonorous laughter and her contagious, sunny smile were the best remedies for my insane anxieties. I am so lucky to have you in my life.

Paris
31 May, 2021

Note to the Reader

The four chapters of this dissertation are self-contained research articles and can be read separately. They are preceded by an introduction which summarizes the research presented in this dissertation. The terms “paper” or “article” are used to refer to chapters. Chapter 2, 3 and 4 are co-authored, which explains the use of the “we” pronoun.

Contents

Introduction **23**

1 Search Frictions in Credit Markets **43**

1 Data . . . 49
1.1 The Credit Register . . . 49
1.2 Loan Rates (M-Contran) . . . 49
1.3 Firm Location, Creation and Geography . . . 50
1.4 Broadband Internet Data . . . 51
2 Search Frictions in Credit Markets . . . 52
2.1 Bank Branch Heterogeneity . . . 52
2.2 Endogenous Firm-Branch Matching . . . 54
2.3 The Geography of Bank Credit . . . 57
2.4 Price Dispersion . . . 58
2.5 Survey Evidence . . . 59
3 Model . . . 60
3.1 Setup . . . 60
3.2 Predictions . . . 64
4 Empirical Context . . . 70
4.1 The Rise of Online Banking . . . 70
4.2 The diffusion of Broadband Internet in France . . . 72
4.3 Identification and Instrumental Variable . . . 73
5 Empirical Approach . . . 75
5.1 Gravity Equation for Aggregate Credit Flows . . . 75
5.2 Firm-Branch Matching . . . 79
6 Results . . . 80
7 Implications for the cost of debt . . . 89
8 Conclusion . . . 91

9 Figures . . . 98
10 Tables . . . 108
A Appendices . . . 111

2 Bank Local Specialization **113**
1 Bank branch networks in France . . . 117
2 Data . . . 120
2.1 Data sources . . . 120
2.2 Data cleaning . . . 121
2.3 Descriptive statistics . . . 122
3 Measuring local bank specialization . . . 123
3.1 Measuring local bank specialization . . . 123
3.2 Stylized facts about local bank specialization . . . 125
4 Empirical strategy . . . 127
4.1 Branch closures and credit supply: firm-bank level analysis . . . 127
4.2 Branch closures and credit supply: firm-level analysis . . . 128
4.3 Branch closures and change in the industry specialization of the bank 128
5 Results . . . 130
5.1 Branch reallocation and credit supply at the firm-bank level . . . 130
5.2 Can multi-bank firms compensate for the effect of a branch reallocation? 131
5.3 The benefits of the local industry specialization of banks . . . 131
6 Conclusion . . . 133
A Appendices . . . 152

3 Aggregate Implications of Credit Relationship Flows **157**
1 Introduction . . . 157
2 Empirical Methodology . . . 162
2.1 Conceptual Foundations: a Flow Approach to Credit Markets . . . 163
2.2 Definitions and Measurement . . . 164
2.3 The French Credit Register . . . 167
2.4 Issues and Adjustments . . . 168
2.5 Summary Statistics and Aggregate Time Series . . . 170
3 Properties of Credit Relationship Flows . . . 171
3.1 Aggregate Patterns . . . 171
3.2 Cyclical Properties . . . 172
3.3 What Drives the Creation and Destruction of Credit Relationships? . . . 173
3.4 Cross-sectional Decomposition at the Relationship Level . . . 174

4 How Do Banks Adjust Their Credit Supply Along Extensive and Intensive Margins? . . . 176
4.1 Simple Credit Decomposition (Decomposition 1) . . . 176
4.2 Secular Trends . . . 177
4.3 Cyclical Fluctuations . . . 177
4.4 Alternative Decomposition: Incumbent vs. New and Severed Credit Relationships and the Importance of the Sub-extensive Margin (Decomposition 2) . . . 179
5 Anatomy of a Credit Crisis and Recovery . . . 182
6 The Extensive Margin Channel of Monetary Policy . . . 183
6.1 Measurement of Monetary Policy Shocks . . . 184
6.2 Aggregate Response . . . 184
6.3 Bank-level Response . . . 186
7 Discussion - Credit Reallocation and Theoretical Implications . . . 187
7.1 Credit Reallocation and Credit Market Fluidity . . . 187
7.2 Implications for Theories of Banking and Credit . . . 189
8 Conclusion . . . 191
A Tables and Figures . . . 196
B Data and Variable Construction . . . 219
B.1 French Credit Register (SCR) . . . 219
B.2 Balance Sheet Data (FIBEN & BRN) . . . 219
B.3 Banking Mergers and Acquisitions (M&As) . . . 220
B.4 Public Banks . . . 220
B.5 Other Reporting Issues . . . 220
C Credit Relationship Flows – Additional Descriptive Results and Robustness Checks . . . 222
D Extensive/Intensive Margin Decompositions – Additional Derivations and Results . . . 230
D.1 Simple Decomposition . . . 230
D.2 Alternative Decomposition 2 . . . 232
D.3 A Third Decomposition: Gross Intensive Credit Flows (Decomposition 3) . . . 233
E Anatomy of a Crisis – Additional Figures . . . 237
F Local Projections – Additional Figures and Results . . . 239
F.1 Conventional vs. Unconventional Monetary Policy . . . 243

4 Technology-induced Trade Shocks? 247

1 Introduction . . . 247
2 Literature . . . 250
3 Data and context . . . 254
3.1 Context: the diffusion of Broadband Internet in France . . . 254
3.2 Data . . . 256
4 Empirical approach . . . 261
4.1 Baseline specification . . . 261
4.2 Validation of the research design . . . 262
5 Baseline Results . . . 264
5.1 Value of imports . . . 264
5.2 Further Robustness checks . . . 267
5.3 Intensive, extensive and sub-extensive margins . . . 270
6 Heterogeneity, mechanisms, and margins of adjustment . . . 272
6.1 Origin-country and type of goods . . . 272
6.2 Analysis of heterogeneous effects at the firm-level . . . 274
6.3 Impact on firm performance and import-intensity . . . 276
7 Conceptual framework for quantification . . . 280
7.1 Baseline model . . . 280
7.2 Extension: exports and labor market clearing wage . . . 284
8 Conclusion . . . 286
A Tables and figures . . . 294
A.1 Baseline tables . . . 294
A.2 Pseudo first-stage . . . 295
B Conceptual framework for quantification: full derivations . . . 296
B.1 Demand and market structure . . . 297
B.2 Firm location choice and city-level sales . . . 297
B.3 Production function . . . 300
B.4 Sufficiency result . . . 301
B.5 Broadband expansion and the sufficiency result . . . 303
B.6 Extension: Exports and endogenous wage . . . 304
C Data appendix . . . 312
C.1 Description of the datasets . . . 312
C.2 List of variables . . . 314
D Additional empirical material . . . 315
D.1 Further robustness checks . . . 315

D.2 Explaining internet coverage : table 319
E Conceptual framework: additional material 328
E.1 Other extensions 328
E.2 Conceptual framework: a simple example 329

Conclusion **333**

Résumé en Français **3**

List of Figures

1.1 Heterogeneity of the effect with respect to distance 84
1.2 Share of credit to remote firms 86
1.3 Share of remote clients 87
1.4 Firm-bank distance 88
1.5 Broadband internet roll-out in France 98
1.6 Branch Rank versus Size (Total credit) 99
1.7 Branch Rank versus Size (#. clients) 100
1.8 Branch Size and Average Distance to Clients 101
1.9 Average Loan Locally and Number of Distant Markets Penetrated 102
1.10 Inter-Submarket and Two-Way Lending 103
1.11 Local Exchanges, Highways and Railroads before 1999 104
1.12 Optimal connection rank predicted vs. observed connection rank 105
1.13 Reduction of the cost of debt triggered by a reduction in search frictions: spatial heterogeneity 106
1.14 Reduction of the cost of debt triggered by a reduction in search frictions: zoom in Paris region 107
1.15 Heterogeneous effect of BI with respect to distance when Paris is connected . 111

2.1 Herfindahl Index at urban unit level 137
2.2 Geographical distribution of bank branches lending to SMEs in France (2016). 138
2.3 The geography of bank branch closures in France, 2010-2017. 139
2.4 Time series of branch closures in France, 2010-2017. 140
2.5 Number of specialized bank branches by borrowing industry. 141
2.6 Share of specialized branches not aligned with the industry specialization of their parent bank, by urban unit (2016). 142
2.7 Impact of a branch closure on credit: firm-bank level analysis. 143
2.8 Impact of a branch closure on credit: firm-level analysis. 144

3.1 The Flow Approach to Credit Markets . . . 201
3.2 Credit Relationships: Concepts and Measurements . . . 202
3.3 Evolution of Banks, Firms, Bank-firm Relationships, and Credit . . . 203
3.4 Share of Credit Relationships by Type and Duration . . . 204
3.5 Credit Relationship Flows . . . 205
3.6 Sources of relationship creation and destruction . . . 206
3.7 First Credit Relationship and Firm Entry . . . 207
3.8 Gross Flows by Credit Size, Type, and Duration . . . 208
3.9 Extensive vs. Intensive Margins: Long-run Trends . . . 209
3.10 Extensive vs. Intensive Margins of Credit – Decomposition 1 . . . 210
3.11 Credit for Incumbents vs. Entering and Exiting Firms . . . 211
3.12 Extensive vs. Intensive Margins of Credit – Decomposition 2 . . . 212
3.13 Anatomy of a Crisis: Unconditional Patterns . . . 213
3.14 Anatomy of a Crisis: Details . . . 214
3.15 Monetary Policy Transmission and Credit . . . 215
3.16 Monetary Policy Transmission and Credit – Easing vs. Tightening . . . 216
3.17 Monetary Policy Transmission and Credit - Bank-level Responses . . . 217
3.18 Monetary Policy Transmission and Credit – Bank Characteristics . . . 218
3.19 Aggregate Credit: French Credit Register (SCR) vs. Flow of Funds . . . 224
3.20 Number of Credit Partners per Bank and per Firm . . . 225
3.21 Trajectories of Credit Growth and Separation Probability . . . 226
3.22 Credit Relationship Flows with 8-Quarter Gaps . . . 227
3.23 Credit Relationship Flows with the 25K Euro Threshold . . . 228
3.24 Aggregate Credit Variations – Cyclical Components (HP Filter) . . . 229
3.25 Creation vs. Destruction Flows: Pre- and Post-2008 . . . 231
3.26 Extensive vs. Intensive Margins: Cyclical Deviations . . . 236
3.27 Anatomy of a Crisis – Decomposition 2 – Creation vs. Destruction . . . 237
3.28 Anatomy of a Crisis – Decomposition 1 . . . 238
3.29 Monetary Policy Shocks – 2002-2018 . . . 239
3.30 Monetary Policy Transmission and Credit – Specification with Lags . . . 240
3.31 Monetary Policy Transmission and Credit – Alternative Monetary Shocks . . . 241
3.32 Monetary Policy Transmission and Credit - ECB Information Shocks . . . 242
3.33 Monetary Policy Transmission – Conventional vs. Unconventional . . . 243
3.34 Monetary Policy Transmission and Credit – micro-level Responses (with Bank-fixed Effects) . . . 244

3.35 Monetary Policy Transmission and Credit – Specification with Credit Decomposition 1 245

4.1 Distribution of $\widetilde{Z}_{it}$: 1999-2007 257
4.2 The progressive roll-out of the DSL technology in France—$\widetilde{Z}$ 258
4.3 Main specification: Log of the value of imports 264
4.4 Counterfactual aggregate trends in overall import 267
4.5 Distribution of Placebo Estimates: Log Imports, β_5 268
4.6 Number of flows and average value per flow 271
4.7 Extensive and sub-extensive margins 271
4.8 Number of products (HS-6) and sourcing countries 272
4.9 $\widehat{\beta}_5$ for different groups of origin-countries 273
4.10 Value of imports by type of goods 274
4.11 Measures of import intensity 277
4.12 Sales and value-added (log) 278
4.13 Value-added per worker (log) 278
4.14 Exports 280
4.15 The evolution of continuous measure of broadband coverage ($\widetilde{Z}_{it}$) around the (discrete) year of the largest increase in $\widetilde{Z}_{it}$ (t_{0i}). 295
4.16 Number of cities by cohort 321
4.17 Distribution of Placebo Estimates: Log Imports, β_4 322
4.18 Distribution of Placebo Estimates: Log Imports 323
4.19 Including multi-city firms in the analysis 324
4.20 Effect on imports by type of goods and origin countries 324
4.21 The progressive roll-out of the DSL technology in Occitanie—$\widetilde{Z}$ 325
4.22 Intermediate inputs over-sales-ratio 325
4.23 Share of foreign inputs (ln), excluding capital goods 326
4.24 Log of the value of imports at firm-level 327
4.25 Long-difference results: 328

List of Tables

1.1 Endogeneous Firm-Branch Matching . . . 56
1.2 Explaining Price Dispersion . . . 59
1.3 Gravity Equation for Inter-regional Credit Flows . . . 81
1.4 Technology-Induced reduction in search frictions . . . 83
1.5 Gravity Equation for Inter-regional Credit Flows . . . 108
1.6 Technology-Induced reduction in search frictions: Pair fixed-effects . . . 109
1.7 Technology-Induced reduction in search frictions with lags . . . 110
1.8 PPML with many zeros in a dynamic setting: simulation results . . . 112

2.1 Firm-bank-level summary statistics. . . . 145
2.2 Firm-level summary statistics . . . 146
2.3 Summary statistics on the specialization of bank branches. . . . 146
2.4 Branch closures and SMEs' access to credit: firm-bank level analysis. . . . 147
2.5 Branch closures and SMES' access to credit: firm-level analysis . . . 148
2.6 Branch closures, branch specialization and SMEs' access to credit: firm-bank level analysis. . . . 149
2.7 Branch closures, branch specialization and SMEs' access to credit: firm-level analysis. . . . 150
2.8 Branch closures, branch specialization and SMEs' access to credit: the role of distance and competition. . . . 151
2.9 Herfindhal Index at urban unit level . . . 152
2.10 What drives bank branch closures ? (1) ANOVA. . . . 152
2.11 What drives bank branch closures ? (2) Linear probability model (department level). . . . 153
2.12 What drives bank branch closures ? Linear probability model (Urban unit level). . . . 154
2.13 Specialization, bank-firm distance and bank branch closures. . . . 155

3.1 Summary Statistics: Aggregate Results 196
3.2 Cyclical Properties of Credit Relationship Flows 197
3.3 Cyclical Properties of Credit Relationship Flows: Cross-sectional Decomposition 198
3.4 Cyclical Properties of Aggregate Variables 199
3.5 Variance Decomposition: Intensive vs. Extensive Margins 200
3.6 Summary Statistics: Cross-sectional Results 222
3.7 Cyclical Properties: Lead-lag Structure 223
3.8 Variance Decomposition: Intensive vs. Extensive Margins (Decomposition 3) 235

4.1 Descriptive statistics at the city-level: 1997-2007 260
4.2 Explaining city broadband coverage: panel analysis 263
4.3 Heterogeneity depending on sector and firm size and productivity 275
4.4 The role of information: Heterogeneity based on origin and products characteristics 276
4.5 Welfare changes for different values of θ and σ 285
4.6 Specification checks for main specification: ln(value of imports) 294
4.7 Parameters and identification 310
4.8 Welfare changes for different values of θ and σ 311
4.9 Specification checks for main specification: asinh(value of imports) 315
4.10 Static panel-fixed effect model 316
4.11 Specification Check: ln(value of imports) , binning $d \in \{-6, -5, -4\}$ 318
4.12 Explaining variation in internet coverage: full panel regressions 320

Introduction

> *"One should hardly have to tell academicians that information is a valuable resource: knowledge is power. And yet it occupies a slum dwelling in the town of economics. Mostly it is ignored."*
>
> – George J. Stigler, *The Economics of Information*

Time has passed since George J. Stigler's seminal article (Stigler, 1961), in which the 1982 Nobel prize winner lamented the lack of research on the topic of *information*. This complaint is now obsolete: in sixty years, the notion of information has penetrated a wide range of research fields in economics. It is now recognized that information is imperfect, obtaining information can be costly and there are important asymmetries of information. The four chapters of this dissertation make an attempt to extend our knowledge on how imperfect information affects both credit and goods markets, by leveraging micro data. I first contextualize each chapter separately. A detailed description of the chapters is then provided.

One early strand of the literature on information has focused on *information acquisition* and developed search models in which individuals incur costs to acquire information (McCall, 1970; Diamond, 1971), thus departing from the abstraction of Walrasian markets in which the price adjusts instantaneously to equilibrate demand and supply. This line of research eventually evolved into the matching-and-bargaining and directed-search models that are now widely used in labor macroeconomics (see Rogerson et al., 2005 for a survey), monetary theory (Kiyotaki and Wright, 1993), financial economics (Duffie et al., 2005; Weill, 2007), urban economics (see Zenou, 2009 for a monograph) and for the analysis of the marriage market (Mortensen, 1988; Shimer and Smith, 2000). An emerging branch of this vast literature has recently highlighted the prominence of search frictions in (international) goods trade (Rauch, 2001; Chaney, 2014; Allen, 2014; Lenoir et al., 2018), acknowledging that it takes time and resources for an exporter to learn about market conditions elsewhere, to find customers

abroad or, symmetrically for an importer to match with the right supplier. Furthermore, this literature has also documented how the diffusion of information and communication technologies (ICTs) reduced such frictions (Jensen, 2007; Aker, 2010; Goyal, 2010; Lendle et al., 2016; Steinwender, 2018; Akerman et al., 2018; Bhuller et al., 2019). Chapter 1 and Chapter 4 of this dissertation directly contribute to this strand of the literature.

The contribution of chapter 1 is to extend the study of search frictions to credit markets. Motivated by empirical evidence I document on local credit markets in France, I propose a theory of firm-bank matching, subject to search frictions. Firms exert significant efforts to locate and match with the right banking partner. Upon matching, agency frictions hinder banks' ability to optimally screen and monitor projects. I structurally estimate my model on French data using the staggered roll-out of Broadband Internet, from 1997 to 2007, as a shock that reduced search frictions. I confirm the model predictions that the allocation of credit was affected by this shock. Finally, I use the structure of my model to quantify the impact of this technology-induced reduction in search frictions on loan prices. I find that broadband internet access reduced the cost of debt for small businesses by 4.9% over the period. In chapter 4, which is joint work with Clément Malgouyres and Thierry Mayer, we study the role of Broadband Internet in reducing search frictions faced by French importers. We document the presence of "technology-induced" trade in France between 1997 and 2007 and assess its impact on consumer welfare. We use the staggered roll-out of broadband internet to estimate its causal effect on the importing behavior of affected firms. Using an event-study design, we find that broadband expansion increases firm-level imports by around 25%. We further find that the "sub-extensive" margin (number of products and sourcing countries per firm) is the main channel of adjustment and that the effect is larger for capital goods.

In parallel with the study of *information acquisition*, a largely separate literature considers environments with limited strategic interactions and has identified the crucial importance of *asymmetric information* (Akerlof, 1970; Spence, 1973; Stiglitz, 1975; Stiglitz and Weiss, 1981), that occurs in transactions where one party has more or better information than the other. In particular, the core ideas about asymmetric information developed in the 1970s continue to play a key role in the analysis of banking and corporate finance. This literature has long highlighted the role of bank-firm relationships in alleviating agency frictions and shaping credit supply (for surveys see, e.g., Boot, 2000; Degryse et al., 2009; Udell, 2015), as banks develop close relationships with borrowers over time which facilitates monitoring and screening. In chapter 2, co-authored with Anne Duquerroy, Jean-Stéphane Mésonnier and Daniel Paravisini, we show that differentiation and specialization may also allow banks to

reduce information asymmetries and gain market power. Using micro-data of the universe of bank-SME relationships in France, we document that banks specialize locally (at the branch level) by industry, and that this specialization shapes the equilibrium amount of borrowing by small firms. For identification, we exploit the reallocation of local clients from closed-down branches to nearby branches of the same bank, which induced quasi-random variation in the match between a firm's industry and the industry of specialization of the lending branch. We show that branch reallocation leads, on average, to a substantial and permanent decline in small firm borrowing. This decline is twice larger for firms whose accounts are reallocated from branches less specialized in their industry than the original one.

Finally, in chapter 3, which is joint work with Yasser Boualam, we study the implication of both agency and search frictions on credit allocation from a macroeconomic standpoint. The common view across most macro-finance models abstracts from the long-term nature of financial contracts and any market frictions that may prevent banks from costlessly forming or severing these credit matches. These models thus downplay the value of relationships and their aggregate consequences and imply that banks can swiftly adjust their number of borrowers in response to shocks. We propose a novel macro perspective on the process of credit intermediation. It aims to provide further empirical evidence on the key and distinctive roles played by both the intensive and extensive margins in shaping aggregate credit fluctuations. Here, we attempt to look behind such fluctuations in order to address first-order questions such as: (i) When aggregate bank credit declines by 5%, is it because the average loan size (i.e., intensive margin) drops by 5%, or is it because 5% of bank-firm matches (i.e., extensive margin) are destroyed? (ii) Does the origin of aggregate credit fluctuations matter? (iii) Do monetary policy shocks impact these margins differently? I now turn to a more detailed description of each chapter.

Chapter 1: Search Frictions in Credit Markets

It is costly for firms to locate and match with the right banking partner, especially for small and medium-sized enterprises (SMEs) who devote time and resources to this search process. Small firms commonly multiply loan applications (2.7 on average) and undergo a time-consuming application process: over 33 hours are spent on loan request paperwork. Overall, about one third of SMEs deplore a difficult and lengthy credit application process. While the effect of informational asymmetries on credit allocation has been well documented (e.g., Akerlof, 1970; Stiglitz and Weiss, 1981; Petersen and Rajan, 1995), little is known about how search frictions affect bank-firm matching and access to credit. Understanding

the role of search frictions is particularly important to policy makers, not only as recent developments in information technology and digitisation in the banking industry are likely to affect search frictions, but also as policies that reduce search costs may differ substantially from policies that reduce traditional agency frictions.

The first contribution of the paper is to develop a theory of firm-bank matching subject to informational frictions. This theory is motivated by new stylized facts I uncover on local credit markets in France: I use a unique data set on small business lending from 1998 to 2005 to document several patterns in the observed loan prices, bank branches heterogeneity and credit flows across cities that suggest, along with survey evidence, the presence of search frictions in credit markets. I then develop a partial equilibrium model of firm-bank matching that is able to capture and explain the observed patterns in the data.[1] The model features two-sided heterogeneity – bank branches and firms – and information frictions of two kinds. First, search frictions hinder firms ability to locate and match with the right financing partner. Second and upon matching, informational asymmetries affect banks ability to screen and monitor projects. By adding structure to the search and matching process, I generate a number of theoretical predictions linking the level of search frictions to (i) the cost of debt for small businesses, (ii) credit flows between cities and (iii) the dynamic of firm-bank matching. In particular, when search costs decrease firms meet with more potential lenders and eventually borrow credit at a lower rate.

The second contribution of the paper is to take the model to the data. I structurally test the main predictions of my model using the staggered diffusion of Broadband Internet (BI) in France, from 1999 to 2007, as a shock on search frictions. I then use my empirical estimates to quantify a technology-induced reduction in the cost of debt for French SMEs, along the lines of my model. The large diffusion of information and communication technologies has indeed represented a profound change for the banking industry and Broadband Internet was the catalyst for this numerical transformation. As digitization proceeded apace, transaction costs decreased and the rise of online financial services allowed firms to search for the best banking partner in a faster and more efficient way, leading to structural changes in banking markets (Hauswald and Marquez, 2003). Kroszner and Strahan (1999) showed how the large adoption of information technology reduced the dependence on geographical proximity between customers and banks, and Petersen and Rajan (2002) documented the erosion of the local nature of small business lending, with increasing distance between small firms and their lenders in the United States but also new communication habits. Similar trends are observed in France: inter-regional credit flows have grown by 15% and the average firm-bank

[1] From a modeling standpoint, my approach is in the spirit of Eaton et al. (2018) and Lenoir et al. (2018).

distance has increased by 10% between 1998 and 2005. My empirical approach provide a causal interpretation for those facts, suggesting that innovations in information technology – namely, Broadband Internet diffusion – have reduced the role of both transaction and search costs in shaping credit outcomes, allowing firms to search for credit further and leading to structural changes in local credit markets.

Identifying the causal effect of new technology adoption on firm-bank matching and credit outcomes is difficult because of its endogeneity. The French data and context allow me to make progress on the causal identification of how technology affects search frictions in credit markets and, in turn, firm-bank matching and credit flows. The third contribution of the paper is that I propose a novel instrument variable strategy for the timing of Broadband Internet expansion, that is based on a theoretical optimal investment plan for infrastructure upgrading. I use a dataset on Broadband Internet availability at the municipality level over the 1999-2007 period compiled by Malgouyres (2017) and combine it with information regarding population density and existing telecommunication infrastructures, namely local copper loops networks and large fiber optic cables. Because Broadband Internet operators relied on these already existing infrastructures to gradually deploy Broadband, I show that the optimal timing of Broadband availability can be predicted using this data, without additional information regarding local economic conditions. This setup provides natural ground for an event-study identifying how information and communication technologies affect credit markets.

My results provide a causal evidence of how search frictions affect firm-bank matching and the allocation of bank credit. First, I find that credit flows between cities follow a gravity equation that is distorted by the staggered roll-out of Broadband Internet. This technology-induced reduction in search frictions triggers an average increase by 6% of the share of credit exchanged between interconnected cities. Consistent with the model's predictions, this effect varies dramatically with the initial level of search frictions: it is higher when two very distant cities are connected. On the contrary, the effect is negative when two neighbouring cities, already economically very closely tied, are connected by internet. For robustness, I provide simulation evidence of the performance of the PPML estimator for the estimation of gravity equations in a dynamic setting and with many zeros, extending the work of Santos Silva and Tenreyro (2006, 2011). I confirm that the Poisson pseudo-maximum likelihood (PPML) estimator with fixed-effects is well behaved and consistently estimates the time-varying treatment effect, event with more than 90% of zeros. I also show that my baseline results is barely not affected by the addition city pair fixed-effects and of the lag dependant variable as covariate.

My main results rely on regressions carried out at the urban unit level, which simply matches the level of the treatment and allows me to deal with an estimation sample of a manageable size (24 million observations). However, some model predictions at a less aggregated level require leveraging bank branch-level data. I further document that Broadband Internet diffusion allows banks to match with new firms located in remote submarkets. Connected banks increase their share of credit lent to firms located outside their city by 10%, and their share of remote clients by almost 12%. As a result, the average distance between a bank and its customers increases by 10% in the medium run after broadband internet access. These results are robust to several potential threats to identification. I find no evidence of pre-expansion differential trends in branch-level outcomes and I show that adding city-level controls do not affect my estimates. Finally, I assess the implications of my findings on the cost of debt for small firms, through the lens of my model. In practice, I plug my empirical estimates into the equation linking search frictions to loan prices. Interpreted within my model, the estimates imply that the reduction in search frictions due to the diffusion of Broadband Internet lowered the cost of debt for small businesses by 4.9% on average. This reduction in the cost of debt displays an interesting spatial heterogeneity. It is stronger in rural areas and medium-sized cities than in the largest french cities. Firms initially located far from bank branches, or that did not have a wide variety of potential banking contacts, benefit more from the reduction of search frictions, since it allows them to match with new or better banking partners. In this respect, the spread of Broadband Internet reduces spatial inequalities in access to credit.

Chapter 2: Local Bank Specialization

In this chapter, co-authored with Anne Duquerroy, Jean-Stéphane Mésonnier and Daniel Paravisini, we investigate how differentiation and specialization allow banks to reduce information asymmetries and gain market power.

Widespread bank branch closures and consolidation in Europe and the United States after the Great Recession have renewed a longstanding policy and academic debate about the nature and implications of bank market power. A vast body of theoretical and empirical work, following Rajan (1992), has explored one source of such market power: the informational monopoly gained through relationship lending. Less studied, and of potentially equal importance, is the market power gained through differentiation and specialization. Lenders that may appear to compete fiercely in an undifferentiated credit market, may in fact enjoy market power in some market segments by tailoring their products and services to particular

clients, industries, or types of financing. Such credit market segmentation may have first order implications for the access to credit by small and opaque bank-dependent firms as well as its cost. Documenting the extent to which banks specialize in a segmented small business credit market, and assessing whether specialization confers market power, poses important data and identification challenges that we address in this paper.

We use unique regulatory data that contains, for the universe of bank-firm relationships in France, the identity and location of the bank branch providing credit. With these data we construct measures of bank and bank-branch industry specialization in narrowly defined geographical credit markets. Figure 2.1 provides stylized motivating evidence for our study. It plots two different measures of credit market concentration by urban unit.[2] The left panel shows the standard concentration measure, calculated using total lending shares. The right panel shows the average of credit concentrations calculated industry-by-industry, which takes into account market segmentation. The difference between these two measures will be larger when the credit market is segmented by industry (e.g., the two measures will be identical if all bank loan portfolios have the same industry composition). Indeed, the fraction of urban units with a very high concentration level (HHI > 0.4) raises from from 21% when measured in the traditional manner, to 49% when measured taking into account credit segmentation by industry (see Table 2.9 for more details). This pattern is consistent with heavy industry segmentation in the small business bank credit market.

The main goal of our empirical analysis is to explore the implications of credit market segmentation for small firms' access to credit. Our working hypothesis is that a firm's credit elasticity of substitution across banks is smaller when specialized banks offer differentiated services. For example, a firm in the construction industry will find more difficult or costly to substitute credit obtained from the bank (or branch) that is specialized in the construction industry than credit obtained from a generalist bank (or branch). A necessary first step is to evaluate the relevant unit of analysis to study specialization. In other words, do banks specialize by industry as a whole? Or is specialization a local, branch-level, phenomenon?

To answer this question we follow the data driven approach developed in Paravisini et al. (2017) to identify banks' sector of specialization using abnormally large portfolio shares. The intuition of the measure is best explained through an example. Suppose 20% of bank credit in an urban area goes to the construction industry and is serviced by five banks. Banks are heterogeneous: while four banks allocate less than 10% of their loan portfolio to construction, the fifth bank allocates more than 40% of its credit portfolio to the sector. This fifth bank

[2] An urban unit is defined as a municipality or a group of municipalities which covers a continuously built up zone (with no more than 200 meters between two constructions) and hosts at least 2,000 inhabitants.

would be identified as a specialist in the construction industry for this urban area. The advantage of using portfolio shares to detect specialist banks is that the identification of the specialization sector is unaffected by the size of the sector or by the market share of each bank in any given location.

Two key stylized facts emerge from this exercise: bank branches tend to specialize by industry, but different branches of the same bank generally exhibit different industrial specializations. More than a third of bank branches in France come out as being specialized in supplying credit to small firms in at least one specific broad industry. Most urban areas include specialized bank branches. Moreover, we observe that most industries exhibit specialized bank branches at the local level. For instance, some 9% of the bank branches present in our sample in 2017 are specialized in funding transportation and storage activities. Overall, this implies that a French SME has a non-negligible probability to get connected to a branch that is specialized in its type of business. When we investigate specialization patterns of the branches within banks, we find that large banks are characterized by a large share of specialized branches (37% for the average bank with more than 10 branches). However, within a bank, different branches tend to be specialized in different industries. In short, industry specialization appears in the data as a widespread but local, branch-level, phenomenon.

Motivated by these stylized facts we turn to measuring the heterogeneity in firm's elasticity of credit substitution by branch specialization. Our empirical research design exploits borrower reallocations across branches due to branch closures. Among bank branches active in SME lending, some 700 branches were closed during our sample period (between 2010 and 2017) throughout the country, due to internal restructuring plans of the main banks' retail activities. Branch closures did not end bank-borrower relationships: all loan accounts in a closing branch were transferred to larger nearby branches of the same bank. Branch reallocation induced variation in the match between the borrower's industry and the industry of specialization of the branch that we exploit to measure the heterogeneity in the elasticity of credit substitution. In the construction firm example above, when branch services are segmented by industry, the transfer of the firm's account to a generalist branch should reduce the equilibrium amount of credit used by the firm, relative to a counterfactual in which the account were transferred to another branch that is also specialized in the construction industry. Branch closures occurred in large waves, and the identity of the closing and absorbing branches were selected by headquarters according to criteria like local bank density, arguably unrelated to the demand for credit of individual firms. The very disaggregated nature of the data also allows using saturated specifications to control for local shocks at the urban unit level, bank shocks, and firm shocks that may occur concurrently with the branch closure.

In the baseline specification we find evidence of a significant drop in the total of credit granted by a bank to a small firm whenever the firm's account is reallocated to a new branch. Including undrawn credit lines, total credit drops by 12% on average over the three years following the effective closing. Part of this decline is substituted with more credit from other banks. However, the average firm's total credit drops permanently by about 4% after an account reallocation, relative to other firms in the same narrow geographical market and industry. We then document the heterogeneity of this decline in equilibrium credit by the match between the borrower's industry and the industry of specialization of the closing and absorbing branches. We find that the magnitude of the decline in credit doubles when a firm's accounts are reallocated from a branch that specializes in its industry to a branch that does not. The magnitude of this estimated effect is robust to controlling for the change in distance associated with the branch closing.

In the cross section, we find that the decline in credit following a branch closure is entirely explained by the variation in industry specialization between branches when the new branch is located in an area characterized by a high level of bank competition. The results are strongly suggestive of a segmented bank credit market, where bank specialization by industry increases the cost of substituting bank sources of financing for small firms.

Chapter 3: Aggregate Implications of Credit Relationship Flows

In this third chapter, co-authored with Yasser Boualam, we study the implication of market imperfections in driving credit allocation from a macroeconomic standpoint. What drives the fluctuations of credit over the business cycle and in the long run? How do banks adjust their credit supply in response to aggregate shocks or policy changes? These questions have been at the forefront of macro-finance and banking research at least since the seminal work of Bernanke (1983). Yet, our understanding of aggregate credit fluctuations and their implications for the real economy remains incomplete on several fronts.

Bank credit is a significant source of financing for the majority of businesses. One particularly important aspect that has been extensively studied at the micro level, yet overlooked in macro, has to do with bank-firm credit relationships. Indeed, a vast theoretical and empirical literature has long highlighted the role of these relationships in terms of alleviating agency frictions and shaping credit supply at the lender-borrower level.[3] It also emphasized the

[3]See Boot (2000) and Degryse et al. (2009) for a survey of earlier work.

existence of cross-sectional heterogeneity in terms of match quality and inherent relationship characteristics such as duration, which can potentially hinder banks' ability to adjust their credit supply in a frictionless way (Boualam (2018)). Conversely, the common view across most macro-finance models either simply assumes homogeneous borrowers and/or lenders, or abstracts from the long-term nature of financial contracts and any market frictions that may prevent banks from costlessly forming or severing these credit matches. These models thus downplay the value of relationships and their aggregate consequences and imply that banks can swiftly adjust the number of their borrowers in response to shocks. They also leave little room for analyzing the process of credit reallocation across bank-firm matches and its dynamics throughout the cycle.

This paper proposes a novel macro perspective on the process of credit intermediation. It aims to provide further empirical evidence on the key and distinctive roles played by both the intensive and extensive margins in shaping aggregate credit fluctuations. Here, we attempt to look behind such fluctuations in order to address first-order questions such as: (i) When aggregate bank credit declines by 5%, is it because the average loan size (i.e., intensive margin) drops by 5%, or is it because 5% of bank-firm matches (i.e., extensive margin) are destroyed? (ii) Does the origin of aggregate credit fluctuations matter? (iii) Do monetary policy shocks impact these margins differently?

To our knowledge, we are the first to show that banks actively adjust both the number *and* the intensity of their relationships, in response to macroeconomic shocks, and that both of these margins represent a significant source of variation in bank lending. These adjustments are somewhat analogous to the ways in which firms constantly adjust both quantity of hours worked and employment, or their capacity utilization and new capital investment.[4] This view may sound intuitive, yet — and surprisingly — a thorough analysis of the dynamics of these margins and their macroeconomic implications remains limited, if not completely absent. Furthermore, we not only establish the quantitative importance of these margins, but we also argue that they are subject to prominently different aggregate behaviors. Thus, disentangling the effects associated with each margin can prove informative about the economic mechanisms at play and the role of credit reallocation, and ultimately yield relevant policy implications.

To shed light on this process, we leverage a key source of information, the French Credit Register, which covers the commercial loan market in France, and is maintained by Banque de

[4]To some extent, our analysis of credit markets follows in the footsteps of Lilien and Hall (1986), who first decomposed the fluctuations in total hours worked into changes in employment and changes in hours worked per employed worker.

France. The data contain granular and nearly exhaustive records of bank-firm matches and corresponding credit exposures over the period 1998-2018. To study the properties of credit relationship flows, we develop an empirical methodology akin to the one pioneered by Davis and Haltiwanger (1992) for labor flows. Our methodology takes into consideration specific characteristics associated with credit market structure and available data. For example, we track data entries for each bank-firm match to determine the time of creation and inferred time of destruction in order to construct the associated gross credit relationship flows. We also account for cross-sectional heterogeneity and the nature of financial contracts through key attributes such as loan size, credit type and maturity, and relationship duration.

Understanding the implications of bank-firm credit relationships is a natural undertaking. However, a dearth of empirical evidence documenting their macro-level properties exists due to the paucity of extensive micro datasets over a sufficiently long period of time. In fact, earlier studies such as Dell'Ariccia and Garibaldi (2005) relied on bank-level call report data. Thus, they cannot identify the involved borrowers and can observe net intensive flows only at the bank level. As a consequence, these studies cannot disentangle extensive from intensive margins, nor precisely capture the underlying magnitude and properties of credit reallocation. Instead, we advance here a novel approach to exploit information available in credit registers, which is typically used in micro settings, to uncover new aggregate findings. Our research establishes the following stylized facts about the extensive and intensive margins of credit:

1. Extensive and intensive margins fluctuate continuously over time. While their persistence is roughly identical, the volatility of the intensive margin is relatively higher.
2. Both margins are important at the business cycle frequency, with the extensive margin contributing about one quarter to one half of the variance in aggregate credit.
3. In the long run, the extensive margin accounts for the bulk of aggregate credit variations.

Our analysis also highlights the following features pertaining to gross credit relationship flows:

1. The creation, destruction, and reallocation of bank-firm relationships coexist throughout the cycle.
2. Creation (inflows) and destruction (outflows) of relationships show greater volatility compared to net flows. Variations in net flows are driven mainly by inflows.
3. Outflows are more volatile for small and short-term loans and credit relationships with duration of less than one year. Inflows are more volatile for relationships with small

loans or lines of credit.

Our results also highlight that credit patterns observed during or in the aftermath of an economic downturn are driven potentially by multiple combinations of extensive and intensive margin sources, suggesting that different economic mechanisms may be at play. A better understanding of the extensive/intensive origin of a credit decline and its bottlenecks can thus be relevant to the design of effective and targeted policy tools. In this context, we analyze how monetary policy gets transmitted through both extensive and intensive margin channels. We show that while the intensive margin responds immediately and strongly to monetary policy surprises, the extensive margin's response is relatively more gradual and subdued during easing regimes. We also note that the extensive margin channel is at play mainly for relatively small banks or those with flexible balance sheets.

Our empirical framework also provides us with tools to better understand the reallocation process occurring in credit markets and the channels through which bank shocks get transmitted to the real economy. In particular, we show that the excess reallocation rate of credit relationships is countercyclical, in line with the cleansing effect of recessions. In addition, yearly (excess) reallocation rates have been steadily declining over the past two decades. These results indicate the existence of factors hampering credit market fluidity and contain relevant theoretical and policy ramifications worthy of further investigation.

Chapter 4: Technology-Induced Trade Shocks ? Evidence from Broadband Expansion in France

In this last chapter, which is joint work with Clément Malgouyres and Thierry Mayer, we study the role of Broadband Internet in reducing search frictions faced by French importers.

From 1995 to 2008, the value of imports by high-income countries has grown twice as fast as global GDP.[5] This acceleration of globalization has induced well-documented labor market impacts (Autor et al., 2016b, summarize the recent literature on the impact of the "China shock" on labor market outcomes), as well as rises in consumer welfare through lower prices and gains in varieties (see Feenstra and Weinstein, 2017, for a recent illustration). This period was also characterized by radical innovations in information and communication technologies (ICT) and by their rapid diffusion throughout the world economy. It is most likely that the "ICT revolution" (Cohen et al., 2004) lowered the cost of carrying out international

[5]The World Bank World Development Indicators report that the ratio of imports over GDP for high-income countries has grown from 43% to 61%.

transactions and contributed to raising the pace of economic integration.[6] To the extent that ICT facilitated international trade, part of the consumer gains induced by trade development should be attributed to the diffusion of ICTs. In this paper, we test this proposition by estimating the effect of the diffusion of broadband internet on the importing behavior of French firms from 1997 to 2007 and by developing a theoretical framework to assess the impact of broadband-induced imports on consumer welfare.

Identifying the causal effect of technology on trade is generically difficult because of its endogeneity. The French data and context allow us to make progress on the causal identification of how technology affects firm-level import behavior. In terms of data, we assemble a novel dataset on broadband internet availability at the municipality level over the 1997-2007 period and combine it with information regarding firms' importing behavior. Regarding the context, we exploit the gradual roll-out of broadband internet in France, which was staggered over several years due to limited funding and completed primarily in order to maximize population coverage with only limited attention paid to local economic conditions. This setup provides natural ground for an event-study identifying how ICT availability affected firms' importing behavior.

We find that the local access to broadband internet leads to a surge in the total value of firm-level imports. Our point estimate implies a 25% increase after five years. When applied to our estimating sample and taking into account dynamic effects, the aggregate effect of broadband expansion on the value of imports in constant dollars over the 1997-2007 period was to increase its growth rate from 75% to 91%, i.e. 16 p.p. or 21%. Our results are robust to several potential threats to identification. First, we find no evidence of pre-expansion differential trends in outcomes. Second, while it is possible that broadband introduction was systematically associated with contemporaneous local economic shocks, we show that adding a rich set of city-level controls for local industry and income dynamics hardly affects our estimates. Additionally, flexibly controlling for changes in local labor market conditions, by including a large set of local fixed effects interacted with year dummies, barely changes our estimates.

We further document changes in importing activities along several margins. We find that the increase in the overall value of imports is primarily associated with an increase in the number of flows and find no effect on the average value per flow (where a flow is defined as the combination of an importing firm, an origin country and a specific product). All types of goods (intermediary, consumption and capital) are affected. We also find that

[6] Development in ICTs, such as internet and cell phones, has also contributed to unifiying domestic markets, in particular in developing economies (Allen, 2014).

broadband internet has a positive impact on firm performance as measured by value-added and sales. Importantly, the import-intensity of firms increases: the ratio of imports over sales is positively affected as well as the share of foreign inputs in overall consumption of intermediates.

Our paper focuses on the importing behavior of firms, an outcome that has been relatively underexplored in comparison to exports (see e.g. Hjort and Poulsen, 2018). This focus is motivated empirically by the fact that imports have been more dynamic than exports over the period and conceptually by the notion that imports matter most directly for consumer. Nevertheless, we present some empirical results on the export side. We find that broadband stimulated exports albeit to a lower extent than import.

Our main results rely on regressions carried out at the city level. This level of aggregation matches the level at which we have variation in exposure to treatment and allows us to capture several margins (intensive margin, extensive margin—firms starting to import—, entry of new firms) which would be missed by a firm-level panel analysis. Moreover, the model we use in the last of the paper delivers naturally a city-level estimable equation for sales and imports. We present firm-level estimates when exploring the heterogeneity of the effect along a number of dimensions, some of which intrinsically defined at the firm-level (size, sector etc.).

In the final part of the paper, we assess the welfare implications of our empirical findings through the lens of a simple but general theoretical model of firm-level imports. We consider a small open economy made of up of several cities. Firms vary idiosyncratically in terms of productivity across cities and choose the cities that maximize their expected profits (Suárez Serrato and Zidar, 2016; Fajgelbaum et al., 2019). We use the standard monopolistic competition cum CES demand setup for final goods so as to link firm-level sales with (quality-adjusted) unit cost. On the production side, hinging upon Blaum et al. (2018), firms combine labor with domestic and imported inputs. Our model features the sufficiency result highlighted by Blaum et al. (2018): the firm-level domestic share of inputs fully characterizes the contribution of imports of intermediates to the reduction of its unit cost. The generality of the setup allows us to remain agnostic regarding which type of trade costs (variable, fixed per destination or product, search friction etc.) is affected by the broadband internet expansion shock. It also generates a very parsimonious framework for welfare analysis. The overall effect of access to fast internet on the consumer price index and the contribution of enhanced access to foreign inputs to that overall effect (which we refer to as the *import channel*) are expressed as a function of two reduced-form estimates and three parameters to be calibrated (either from descriptive statistics or from the existing relevant literature).

Under our preferred values for calibrated parameters, our event-study estimates imply that broadband internet led to a price index decrease of 1.85%. The import channel contributed up to 0.75%, i.e. about 40% of the overall effect. We complete this analysis with an extension that considers the impact of broadband on exports and nominal wages. Results from the extended welfare analysis indicate that the *export channel* played a much more limited role than the import one in driving welfare gains from broadband expansion.

References

Aker, J. C. (2010). Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics*, 2(3):46–59.

Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Akerman, A., Leuven, E., and Mogstad, M. (2018). Information frictions, internet and the relationship between distance and trade. *mimeo*, page 48.

Allen, T. (2014a). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Allen, T. (2014b). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Autor, D. H., Dorn, D., and Hanson, G. H. (2016). The china shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics*, 8:205–240.

Bernanke, B. S. (1983). Non-monetary effects of the financial crisis in the propagation of the great depression. Technical report, NBER Working Paper.

Bhuller, M., Kostøl, A., and Vigtel, T. C. (2019). How Broadband Internet Affects Labor Market Matching. Memorandum 10/2019, Oslo University, Department of Economics.

Blaum, J., Lelarge, C., and Peters, M. (2018). The gains from input trade with heterogeneous importers. *American Economic Journal: Macroeconomics*, 10(4):77–127.

Boot, A. W. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, 9(1):7–25.

Boualam, Y. M. (2018). Credit markets and relationship capital. *Working Paper at University of Pennsylvania,*.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634.

Cohen, D., Garibaldi, P., and Scarpetta, S. (2004). *The ICT revolution: Productivity differences and the digital divide*. Oxford University Press.

Davis, S. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, 107(3):819–863.

Degryse, H., Kim, M., and Ongena, S. (2009). *Microeconometrics of Banking Methods, Applications, and Results*. Oxford University Press.

Dell'Ariccia, G. and Garibaldi, P. (2005). Gross credit flows. *The Review of Economic Studies*, 72(3):665–685.

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory*, 3(2):156–168.

Duffie, D., Gârleanu, N., and Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, 73(6):1815–1847.

Eaton, J., Kortum, S., and Kramarz, F. (2018). Firm-to-Firm Trade: Imports, exports, and the labor market. Discussion papers 16048, Research Institute of Economy, Trade and Industry (RIETI).

Fajgelbaum, P. D., Morales, E., Suárez Serrato, J. C., and Zidar, O. (2019). State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376.

Feenstra, R. C. and Weinstein, D. E. (2017). Globalization, markups, and us welfare. *Journal of Political Economy*, 125(4):1040–1074.

Goyal, A. (2010). Information, direct access to farmers, and rural market performance in central india. *American Economic Journal: Applied Economics*, 2(3):22–45.

Hauswald, R. and Marquez, R. (2003). Information Technology and Financial Services Competition. *The Review of Financial Studies*, 16(3):921–948.

Hjort, J. and Poulsen, J. (2018). The arrival of fast internet and employment in africa. *American Economic Review*.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The Quarterly Journal of Economics*, 122(3):879–924.

Kiyotaki, N. and Wright, R. (1993). A search-theoretic approach to monetary economics. *The American Economic Review*, 83(1):63–77.

Kroszner, R. S. and Strahan, P. E. (1999). What drives deregulation? economics and politics of the relaxation of bank branching restrictions. *The Quarterly Journal of Economics*, 114(4):1437–1467.

Lendle, A., Olarreaga, M., Schropp, S., and Vézina, P.-L. (2016). There goes gravity: ebay and the death of distance. *Economic Journal*, 126(591):406–441.

Lenoir, C., Mejean, I., and Martin, J. (2018). Search Frictions in International Good Markets. Technical report.

Lilien, D. and Hall, R. (1986). Cyclical fluctuations in the labor market. *Handbook of Labor Economics*, 2(Part C):1001–1035.

Malgouyres, C. (2017). The impact of chinese import competition on the local structure of employment and wages: Evidence from france. *Journal of Regional Science*, 57(3):411–441.

McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126.

Mortensen, D. T. (1988). Matching: Finding a partner for life or otherwise. *American Journal of Sociology*, 94:S215–S240.

Paravisini, D., Rappoport, V., and Schnabl, P. (2017). Specialization in Bank Lending: Evidence from Exporting Firms. (12156).

Petersen, M. and Rajan, R. (2002). Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570.

Petersen, M. A. and Rajan, R. (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of finance*, 47(4):1367–1400.

Rauch, J. E. (2001). Business and social networks in international trade. *Journal of Economic Literature*, 39(4):1177–1203.

Rogerson, R., Shimer, R., and Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988.

Santos Silva, J. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Santos Silva, J. and Tenreyro, S. (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112(2):220–222.

Shimer, R. and Smith, L. (2000). Assortative matching and search. *Econometrica*, 68(2):343–370.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Steinwender, C. (2018). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–96.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Stiglitz, J. (1975). The theory of "screening," education, and the distribution of income. *American Economic Review*, 65(3):283–300.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410.

Suárez Serrato, J. C. and Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, 106(9):2582–2624.

Udell, G. F. (2015). SME Access to Intermediated Credit: What Do We Know and What Don't We Know? In Moore, A. and Simon, J., editors, *Small Business Conditions and Finance*, RBA Annual Conference Volume (Discontinued). Reserve Bank of Australia.

Weill, P.-O. (2007). Leaning against the wind. *The Review of Economic Studies*, 74(4):1329–1354.

Zenou, Y. (2009). *Urban Labor Economics*. Cambridge University Press.

1

Search Frictions in Credit Markets

"Le véritable voyage de découverte ne consiste pas à chercher de nouveaux paysages, mais à avoir de nouveaux yeux."

– Marcel Proust, *La Recherche du temps perdu*

Abstract

Motivated by empirical evidence I uncover on local credit markets in France, I propose a theory of firm-bank matching, subject to search frictions. Firms undergo a costly search process to locate and match with the right banking partner. Upon matching, agency frictions hinder banks' ability to optimally screen and monitor projects. Structurally, I estimate my model on French data using the staggered roll-out of Broadband Internet, from 1997 to 2007, as a shock to search frictions. I confirm the model predictions that the allocation of credit was affected by the shock. Finally, I use the structure of my model to quantify the impact of this technology-induced reduction in search frictions on loan prices. I find that broadband internet access reduced the cost of debt for small businesses by 4.9%.

JEL classification: L22, L23, D83

Keywords: Search frictions, Broadband internet, Firm-bank matching.

Introduction

It is costly for firms to locate and match with the right banking partner, especially for small and medium-sized enterprises (SMEs) who devote time and resources to this search process. Small firms commonly multiply loan applications (2.7 on average) and undergo a time-consuming application process: over 33 hours are spent on loan request paperwork. Overall, about one third of SMEs deplore a difficult and lengthy credit application process. While the effect of informational asymmetries on credit allocation has been well documented (e.g., Akerlof, 1970; Stiglitz and Weiss, 1981; Petersen and Rajan, 1995), little

is known about how search frictions affect bank-firm matching and access to credit. Understanding the role of search frictions is particularly important to policy makers, not only as recent developments in information technology and digitisation in the banking industry are likely to affect search frictions, but also as policies that reduce search costs may differ substantially from policies that reduce traditional agency frictions.

The first contribution of the paper is to develop a theory of firm-bank matching subject to informational frictions. This theory is motivated by new stylized facts I uncover on local credit markets in France: I use a unique data set on small business lending from 1998 to 2005 to document several patterns in the observed loan prices, bank branches heterogeneity and credit flows across cities that suggest, along with survey evidence, the presence of search frictions in credit markets. I then develop a partial equilibrium model of firm-bank matching that is able to capture and explain the observed patterns in the data.[1] The model features two-sided heterogeneity – bank branches and firms – and information frictions of two kinds. First, search frictions hinder firms' ability to locate and match with the right financing partner. Second and upon matching, informational asymmetries affect banks' ability to screen and monitor projects. By adding structure to the search and matching process, I generate a number of theoretical predictions linking the level of search frictions to (i) the cost of debt for small businesses, (ii) credit flows between cities and (iii) the dynamic of firm-bank matching. In particular, when search costs decrease, firms meet with more potential lenders and eventually borrow credit at a lower rate.

The second contribution of the paper is to take the model to the data. I structurally test the main predictions of my model using the staggered diffusion of Broadband Internet (BI) in France, from 1999 to 2007, as a shock on search frictions. I then use my empirical estimates to quantify a technology-induced reduction in the cost of debt for French SMEs, along the lines of my model. The large diffusion of information and communication technologies has indeed represented a profound change for the banking industry and Broadband Internet was the catalyst for this numerical transformation. As digitization proceeded apace, transaction costs decreased and the rise of online financial services allowed firms to search for the best banking partner in a faster and more efficient way, leading to structural changes in banking markets (Hauswald and Marquez, 2003). Kroszner and Strahan (1999) showed how the large-scale adoption of information technology reduced the dependence on geographical proximity between customers and banks, and Petersen and Rajan (2002) documented the erosion of the local nature of small business lending, with increasing distance between small firms

[1] From a modeling standpoint, my approach is in the spirit of Eaton et al. (2018) and Lenoir et al. (2018).

and their lenders in the United States and new communication habits. Similar trends are observed in France: inter-regional credit flows have grown by 15% and the average firm-bank distance has increased by 10% between 1998 and 2005. My empirical approach provide a causal interpretation for those facts, suggesting that innovations in information technology – namely, Broadband Internet diffusion – have reduced the role of both transaction and search costs in shaping credit outcomes, allowing firms to search for credit further and leading to structural changes in local credit markets.

Identifying the causal effect of new technology adoption on firm-bank matching and credit outcomes is difficult because of its endogeneity. The French data and context allow me to make progress on the causal identification of how technology affects search frictions in credit markets and, in turn, firm-bank matching and credit flows. The third contribution of the paper is that I propose a novel instrumental variable strategy for the timing of Broadband Internet expansion, that is based on a theoretical optimal investment plan for infrastructure upgrading. I use a dataset on Broadband Internet availability at the municipality level over the 1999-2007 period compiled by Malgouyres (2017) and combine it with information regarding population density and existing telecommunication infrastructures, namely local copper loops networks and large fiber optic cables. Because Broadband Internet operators relied on these already existing infrastructures to gradually deploy Broadband, I show that the optimal timing of Broadband availability can be predicted using this data, without additional information regarding local economic conditions. This setup provides natural ground for an event-study identifying how information and communication technologies affect credit markets.

My results provide causal evidence of how search frictions affect firm-bank matching and the allocation of bank credit. First, I find that credit flows between cities follow a gravity equation that is distorted by the staggered roll-out of Broadband Internet. This technology-induced reduction in search frictions triggers an average increase by 6% of the share of credit exchanged between interconnected cities. Consistent with the model's predictions, this effect varies dramatically with the initial level of search frictions: it is higher when two very distant cities are connected. On the contrary, the effect is negative when two neighbouring cities, already economically very closely tied, are connected by internet. For robustness, I provide simulation evidence of the performance of the PPML estimator for the estimation of gravity equations in a dynamic setting and with many zeros, extending the work of Santos Silva and Tenreyro (2006, 2011). I confirm that the Poisson pseudo-maximum likelihood (PPML) estimator with fixed-effects is well behaved and consistently

estimates the time-varying treatment effect, even with more than 90% of zeros. I also show that my baseline results are barely affected by the addition city pair fixed-effects and of the lag dependent variable as covariate.

My main results rely on regressions carried out at the urban unit level, which simply matches the level of the treatment and allows me to deal with an estimation sample of a manageable size (24 million observations). However, some model predictions at a less aggregated level require leveraging bank branch-level data. I further document that Broadband Internet diffusion allows banks to match with new firms located in remote submarkets. Connected banks increase their share of credit lent to firms located outside their city by 10%, and their share of remote clients by almost 12%. As a result, the average distance between a bank and its customers increases by 10% in the medium term after broadband internet access. These results are robust to several potential threats to identification. I find no evidence of pre-expansion differential trends in branch-level outcomes and I show that adding city-level controls do not affect my estimates.

Finally, I assess the implications of my findings on the cost of debt for small firms, through the lens of my model. In practice, I plug my empirical estimates into the equation linking search frictions to loan prices. Interpreted within my model, the estimates imply that the reduction in search frictions due to the diffusion of Broadband Internet lowered the cost of debt for small businesses by 4.9% on average. This reduction in the cost of debt displays an interesting spatial heterogeneity. It is stronger in rural areas and medium-sized cities than in the largest French cities. Firms initially located far from bank branches, or which did not have a wide variety of potential banking contacts, benefited more from the reduction of search frictions, since it allows them to match with new or better banking partners. In this respect, the spread of Broadband Internet reduces spatial inequalities in access to credit.

Related literature. This paper contributes to the literature on corporate finance and search and matching.

There is a vast theoretical and empirical literature in corporate finance and microeconomics on the role that informational frictions plays in hampering firms' access to credit, starting with seminal papers by Akerlof (1970) and Stiglitz and Weiss (1981).[2] This literature has long highlighted the role of bank-firm relationships in alleviating agency frictions and shaping

[2]see also Sharpe (1990); Berger and Udell (1995); Rajan (1992); Berger and Udell (2002); Agarwal and Hauswald (2010); Ioannidou and Ongena (2010); Degryse and Ongena (2005); Drexler and Schoar (2014); Nguyen (2019)

credit supply (for surveys see, e.g., Boot, 2000; Degryse et al., 2009; Udell, 2015). Yet, the role of search frictions, in particular in the formation of bank-firm relationships, has been overlooked despite its potentially equal importance.[3] I propose a novel theory of firm-bank matching and SME access to credit that formally introduces search and contracting frictions. I first provide empirical evidence that emphasizes the importance of search and transaction costs in local credit markets. Then, I write a partial equilibrium model of firm-bank matching subject to both search and agency frictions that I structurally take to the data.[4] As far as I know, this paper is the first one to account for search frictions in the formation of bank-firm relationships from a micro perspective and provide causal evidence of how a reduction in search frictions affects the allocation of credit.[5] This contribution has important policy implications, as an environment characterized by search frictions may not only amplify and propagate shocks (den Haan et al., 2003; Wasmer and Weil, 2004) but also generate slow recoveries (Boualam, 2018).

This paper is also directly related to the literature that explores the determinants of bank-firm matching. Firm and bank size appear as key characteristics (Stein, 2002; Hubbard et al., 2002; Cole et al., 2004; Berger et al., 2005), along with geographic proximity (Petersen and Rajan, 1995, 2002), export country specialization (Paravisini et al., 2015), monitoring capacity (Jing, 2014) and bank capitalization (Schwert, 2018). However, little evidence exists on the importance of branch characteristics, despite the fact that bank branches and loan officers are the main contact point for SMEs searching for the right banking partner (Berger et al., 1997). I contribute to this literature by leveraging bank branch-level data. I show that branch characteristics (size, distance to client, growth, specialization) matter for the matching with small firms. I use this empirical evidence as a motivation for my theoretical framework, where bank branches are heterogeneous in size and distance to clients.

My model delivers a gravity equation for inter-regional credit flows which echoes a nascent literature in finance and macroeconomics that studies gravity equations for cross-border

[3]Macro-finance papers have studied the aggregate effect of search frictions in credit market, starting with den Haan et al. (2003) and Wasmer and Weil (2004). Recently, Argyle et al. (2019) and Allen et al. (2019) have shown how search frictions affect mortgage and consumer credit markets.

[4]I model loan prices are a function of agency frictions that depend on productivity and distance and that capture screening and monitoring costs (Degryse and Ongena, 2005; Hauswald and Marquez, 2006).

[5]Recently Boualam (2018), that builds a general equilibrium theory of bank lending relationships in an economy subject to search frictions and limited enforceability, in a related macro approach. The interaction between search and agency frictions generates a slow accumulation of lending relationship capital, distorts the optimal allocation of credit and leads to slow recoveries. While its approach features a directed search process, I model search as random. Not only this is for tractability purpose, but it is also consistent with the fact that entrepreneurs – who can't devote much time and ressources to this costly search – don't observe bank branches and loan officers characteristics before meeting with them.

equity flows (Portes and Rey, 2005), bonds and bank holdings (Coeurdacier and Martin, 2009) and cross-border asset holdings (Okawa and van Wincoop, 2012). Recently, Brei and von Peter (2018) estimated gravity equations on international banking flows using a PPML approach. They find a substantial role of distance for banking, even with immaterial transport cost, pointing to the role of information frictions. My paper complements and extends these findings by estimating a gravity equation for within-country bank credit flows with two types of informational asymmetries. It confirms the prominent role of distance, even for local credit markets, but also underlines the relevance of search frictions.

This paper also relates to the vast literature in labor (see for a survey, e.g., Rogerson et al., 2005) and trade (Rauch, 2001; Chaney, 2014; Allen, 2014; Lenoir et al., 2018) that studies how search and matching frictions affect firms' ability to produce. This literature has long highlighted that it takes time and resources for a worker to land a job, for a firm to fill a vacancy, for an exporter to find customers abroad or, symmetrically, for an importer to match with the right supplier remotely. I contribute to this literature by showing that search and matching frictions also affect firms' ability to raise external finance. While the theoretical approach developed herein is in the spirit of trade models (Eaton et al., 2018; Lenoir et al., 2018) where firms undergo a random search process, it incorporates agency frictions that replace traditional iceberg transport costs. Finally, some recent papers explore how the diffusion of information and communication technologies – including Broadband Internet – affect such frictions (Allen, 2014; Lendle et al., 2016; Steinwender, 2018; Akerman et al., 2018; Magouyres et al., 2019; Bhuller et al., 2019). My contribution to this literature is twofold. First, I structurally estimate the impact of a technology-induced reduction in search frictions on credit markets and credit search. I provide a causal interpretation of recent structural changes observed in banking markets (Kroszner and Strahan, 1999; Petersen and Rajan, 2002) such as the erosion of the dependence on geographical proximity between customers and banks and the change in competition, eroding banks' rents, in line with Hauswald and Marquez (2003). Second, I develop a novel IV for the timing of Broadband Internet diffusion. As the observed pace of expansion of Broadband Internet may be endogenous, I show how to generate a connection timing that only depends on such ex-ante city-level characteristics as population density and distance to infrastructure.

The rest of the paper is organized as follows. Section 1 presents the data. Section 2 documents new stylized facts on local credit markets in France that set the ground and serve as foundation for our theoretical framework. Section 3 presents the model and its main predictions that will guide the empirical analysis. Section 4 presents the empirical context

and details the instrument variable strategy. Section 5 describes my empirical methodology. Section 6 details the results. Finally, Section 7 studies the implications for the cost of debt through the lens of my model. Section 8 concludes.

1. Data

In this section, I provide a brief description of the data I use to study local credit markets in France and how they were affected by Broadband Internet diffusion over the period 1998-2005. I combine multiple proprietary data sets from the Banque de France about firm-bank relationships, branch-level credit exposure, interest rates of new loans, with unique data on Broadband Internet expansion at the city-level.

1.1. The Credit Register

The French Credit Register is a large data set of bank-firm linkages available at the Bank of France over the period 1998-2005. This credit register aims to collect data on bank exposures to residents on a monthly basis to monitor and control systemic risk. The monthly data comes from reports by credit institutions that are mandatory provided that their commitments or risk exposures on a company as defined by a legal unit and referenced by a national identification number (SIREN), reach a total of EUR 75,000. I use a yearly version of the credit register for convenience as I already deal with an oversized dataset. Monthly reports encompass the funds made available or drawn credits, the bank's commitments on credit lines and guarantee as well as medium and long-term leasing, factoring and securitized loans at the branch level. Recipients are single businesses, corporations, sole-proprietorship engaged in professional activities. They may be registered in France or abroad. Reporting financial intermediaries include all resident credit institutions, investment firms, and other public institutions. In 2005, this raw data set excluding individual entrepreneurs covers information on more than 1.9 millions of bank-firms relationships, corresponding to more than 1.4 million of unique firms or corporations (SIREN), 341 unique banks and 15,925 bank branches serving firms. The smallest banks in my sample own only one branch. In contrast, the largest one (i.e. Crédit Lyonnais) owns more than 2,800 branches all across the country.

1.2. Loan Rates (M-Contran)

The M-Contran database gathers quarterly survey about all new loans granted to French firms from 2006 Q1 to 2016 Q4. This information is collected by the Banque de France in

order to compute quarterly aggregate statistics on the interest rates of new loan contracts, with breakdowns by types of loans, borrowing sectors and types of credit institution. It also enables Banque de France to estimate and publish usury interest rates, an upward limit on lending rates set by French law. All main credit institutions report exhaustive information for all new individual loans from their reporting branches granted during the first month of each quarter. The initial dataset reports, on average, about 100,000 new loans in each quarter. In addition to interest rates,the survey provides rich information on a wide range of relevant loans characteristics, such as the size of the loan, the loan's precise purpose (investment, treasury, leasing etc.), its maturity at issuance, whether it is fixed-rate or adjustable rate, and whether it is secured or not.

1.3. Firm Location, Creation and Geography

Firm Location and Creation. In order to conduct different empirical exercises along this paper, I gather a rich set of (i) administrative data on the creation date and the location French firms' establishments (from the exhaustive SIRENE database). Firms' establishment locations allow me to identify mono-establishment or mono-city firms (which is a key feature of my identification strategy), the year of creation is useful to study firm entry and first banking partner choice in section 2.2.

Geographical data. I gather geographical data for mainland France about more than 36.000 municipalities, 2,000 urban units and 762 urban areas. An urban unit (UU) is a commune or group of communes that includes on its territory a built-up area of at least 2,000 inhabitants where no dwelling is separated from the nearest one by more than 200 meters. In addition, each municipality concerned has more than half of its population in this built-up area. The largest geographical unit that we consider in this paper is the urban area (UA). An urban area is defined as a group of municipalities, all in one piece and without enclaves, consisting of an urban pole with more than 10,000 jobs, and rural municipalities or urban units where at least 40% of the resident population with a job works in the pole or in municipalities attracted by it. Finally, municipalities or cities are the finest unit of measurement that I use for distance computation, firm and branch locations. For most of my empirical analysis, I aggregate bank branches credit exposure, distance to clients, etc., at the urban unit level. I then use urban unit or city to refer to this geographical unit.

1.4. Broadband Internet Data

I use the unique data collected by Magouyres et al. (2019) about the date of upgrade to ADSL for each Local Exchange (LE)'s in mainland France. The historical operator (France Télécom) had to make this data available to other operators as well as websites allowing consumers to gauge the quality of their line for regulatory reasons. Additionally, the authors gather data from the regulatory agency (ARCEP) regarding the geographical coverage of each LE. Combining both datasets, Magouyres et al. (2019) construct a continuous measure[6] of broadband access of city i at year t, denoted $Z_{it}^{\text{city-level}}$, which is a time-weighted percentage of area covered in city i. I construct a additional continuous measure of connectivity at the urban-unit level, $Z_{ut}^{\text{urban-unit}}$ which is simply the (weighted) sum of $Z_{it}^{\text{city-level}}$ for each $i \in u$:

$$Z_{ut}^{\text{urban-unit}} = \sum_{i \in u} w_i \cdot Z_{it}^{\text{city-level}} \tag{1.2}$$

where $i \in u$ denotes the municipalities included in urban unit u and w_i is the weight of city i in the total population of urban unit u. Equation 1.2 implies that $Z_{ut}^{\text{urban-unit}}$ is continuous between 0 and 1. Namely, 0 implies that no firm located in u enjoy an ADSL connection during year t, while 1 indicates that all firms benefit from it all the year long. Figure 1.5 shows the roll-out of Broadband Internet for all urban-units in mainland France, from 1998 to 2005. The dark areas represent a large degree of coverage (a high Z_{ut}).

[FIGURE 1.5 ABOUT HERE]

In 2000, those are confined to the few major cities of France, surrounded by a large majority of no-ADSL territories. By 2003, the treatment has largely spread to lower scale-municipalities, although large parts of France remain dependent on the old technology.

[6] Formally, the measure used by Magouyres et al. (2019) writes:

$$Z_{it} = \sum_{b \in i} \frac{\#\text{ days with access in } b \text{ since Jan 1st of } t}{\#\text{days in year } t} \times \frac{\text{area}_{bt}}{\sum_{b \in i} \text{area}_{bt}} \tag{1.1}$$

$\widetilde{Z}_{it}$ will be equal to one if all of its areas have had access for the entire year. It will be equal to 1/2 if the entire city has had access to broadband over half the year t.

2. Search Frictions in Credit Markets

In this section, I present a body of novel facts about French credit markets that suggest the presence of search frictions: I first document bank branch heterogeneity and endogenous firm-branch matching; then, I describe the geography of credit flows and I provide evidence of substantial price dispersion; Finally, I mirror those empirical facts with new survey evidence.

2.1. Bank Branch Heterogeneity

In a highly competitive and decentralized banking sector, large national banks compete locally through their branch networks, across multiple geographic submarkets. Local bank branches and loan officers are therefore the main contact point for entrepreneurs searching for the right banking partner and branch offices characteristics appear as critical factors to firms, especially SMEs, when choosing their financial services providers (Berger et al., 1997). While prior literature shows that the matching of firms and banks is endogeneous and depends on firm and bank size (Stein, 2002; Hubbard et al., 2002; Cole et al., 2004; Berger et al., 2005), geographic proximity (Petersen and Rajan, 1995, 2002), export country specialization (Paravisini et al., 2015), monitoring capacity (Jing, 2014) and bank capitalization (Schwert, 2018), little evidence exists on the importance of branch characteristics. In this section, I focus on bank branches – rather than bank – heterogeneity. I document four important facts. Branches differ markedly from each others with respect to (i) their total credit exposure, (ii) their average distance to clients (iii) the number of markets in which they operate and (iv) their portfolio specialization.

Branch Size. Figure 1.6 and 1.7 display the distribution of branch size for the last quarter of 2005. I compute the size of a branch as its total credit exposure (1.6) and, alternatively, its number of clients (1.7). Then, I rank branches by size[7]: #1 being the largest branch, #2 the second largest, and so on; finally, I plot the log (Size) versus the log (Rank).

[FIGURES 1.6 AND 1.7 ABOUT HERE]

[7] Regressing the log rank on log size, I find the following:

$$\log\text{ (Rank)} = 12.29 \underset{[0.002]}{\mathbf{-1.04}} \cdot \log\text{ (Size)}$$

The relationship is close to a straight line (R2=0.95), and the slope is very close to 1 (the standard deviation of the estimated slope is 0.02). This means that the rank of a bank branch is essentially proportional to the inverse of its size. A slope of approximately 1 has been found repeatedly using data on city and firm size, stock markets returns, etc. (Gabaix, 2016)

The relationship between log size and log rank is close to a straight line and the slope is very close to 1. This indicates that the distribution of branch size follows a power law (i.e. Pareto distribution): a few number of very large branches grant credit to many firms ($\geq$ 10.000 clients) while a vast majority of small offices only finance 10 to 20 clients.

Branch-Firm Distance. Figure 1.8 shows the positive correlation between branch size (measured as total credit exposure and, alternatively, as the number of clients) and average square geographic distance between the branch and its clients, in kilometers, for the last quarter of 2005.

[FIGURE 1.8 ABOUT HERE]

The average squared distance from a branch to its clients is a power function of the branch size: the bigger the branch the larger the geographic distance is. Bank branches differ in their role and their ability to finance remote clients. Smaller branches focus on their very local market and concentrate on proximity lending: example (branches with less than 10 clients are in average 15.5 km from them). On the contrary, largest branches grant credit to firms located in other cities or regions (branches with more than 100 clients (top 1%) are, in average, 80 km from them). This heterogeneity may be the result of banks internal organization and simply reflects the hierarchy between local branch offices and large business centers. This may also be driven by branch productivity (larger branches being able to screen and monitor more efficiently remote clients), specialization or local economic conditions, but this question remains beyond the scope of this article.

Branch Size and Market Entry. How does a bank branch ability to finance distant clients relate to its local activity? I rank and group bank branches based on the number of remote submarkets (i.e. urban units) where they operate: all my branches lend at least to one submarket while none are active in all. Figure 1.9 depicts average credit size in branch local submarket ("at home") for the group of branches that operate at least in k remote submarkets, with k on the x axis.

[FIGURE 1.9 ABOUT HERE]

The relationship between the average credit exposure in the branch local submarket and the number of remote submarkets penetrated is strongly positive and corroborates the findings

of 1.8. The larger a branch is locally, the higher the number of distant submarkets served. This striking regularity reveal another facet of the bank branch heterogeneity that relates first to the likelihood, for a firm, of meeting with a bank branch located outside its local submarket and, second, to the ability of a loan officer to remotely screen and monitor this firm.

Bank Branch Specialization. Using similar data and methodology than Duquerroy et al. (2019), I show that, within a bank and a urban unit, branches specialize in several dimensions: industry, type of loans and type of businesses. In particular, some bank branches appear to finance heavily SMEs in comparison with other branches located in a same submarket and belonging to a same banking network. I define a branch to be specialized in an industry (respectively, size category) if its portfolio share of lending to firms in an industry (respectively, a size category) is a right-tail outlier in the distribution of portfolio shares of lending by all branches within the same urban unit.

2.2. Endogenous Firm-Branch Matching

In section 2.1, I documented the fact that bank branches markedly differ from each others in several dimensions. Thus, search frictions may arise if firms pay attention to those branch characteristics and devote time and ressources to locate and match with the *ideal* interlocutor. Does this observed heterogeneity affect how firms search for the right banking partner ? Does it lead to endogenous matching ? To address those questions, I investigate the importance of branch characteristics in new firms matching decisions.

Data Sources. I use a new dataset of firm creations from 2002 to 2005 for mono-establishment firms that I combine with information from the Credit Register about firms first realized banking match. The number of firm creations in my sample ranges from 9,958 in 2002 to 14,178 in 2005. After their establishment, firms match with a single bank branch: the period of time running from firm entry to firm first banking match is two years on average. However, more than 50% of those new firms find their banking partner within their first year of existence. For the sake of simplicity, I keep in my final sample firms that match with their first banking partner in less than 5 years, which represent 90% of the observations: I end up with 45,685 new firms entering 711 distinct submarkets and matching with 12,952 bank branches. For each of those new firms, I build the universe of possible banking partners, which I define as all the existing branches located in the same urban area that are active the year of entry. Over all possible pairs, I observe only a single realized match.

Regression Specification. I present reduced-form evidence of endogenous firm-branch matching: a new firm i, when entering a local credit submarket, is more likely to match with a bank branch j that ex-ante shares some common specialisation and size characteristics with the firm. In particular, conditional on distance, a newly created SME from sector s is more likely to borrow from a branch that specializes in lending to small firms or to firms belonging to the sector s. Additionally, I find that larger and growing branches appear to easily attract and match with new clients. Formally, I estimate a Probit regression of different specifications of the following equation :

$$\begin{aligned}\mathbb{P}(\text{ Match}_{i,j,t} \mid \text{observables }) = \Phi\Big(&\alpha \cdot \text{Distance}_{i,j} \\ &+ \beta_1 \cdot \text{Size}_{j,t-1} + \beta_2 \cdot \text{Size Growth}_{j,t-1} \\ &+ \beta_3 \cdot \text{Spec. Sector}_{j,t-1} + \beta_4 \cdot \text{Spec. Size}_{j,t-1} \\ &+ \text{Controls}_{i,j,u,t}\Big)\end{aligned} \tag{1.3}$$

where Φ is the c.d.f. of the standard normal distribution; Spec. Sector$_{j,t-1}$ (respectively Spec. Size$_{j,t-1}$) takes the value 1 if branch j is specialized in lending to firms from the same sector (respectively same size) as firm i, the year before firm i entry. Size$_{j,t-1}$ is the log size of branch j, measured as total credit exposure and, alternatively, as the number of clients and Size Growth$_{j,t-1}$ is the growth rate of branch j, both measured the year before firm i entry. Distance$_{i,j}$ is simply the log geodesic distance in kilometers between branch j and firm i. All explanatory variables are lagged so the branch j characteristics are measured before firm i entry and are not contaminated by the realized firm-branch matching. The dependent variable is an indicator function that takes the value 1 if firm i is borrowing from branch j at t.

Coefficients Interpretation. The coefficient α controls for the direct impact of proximity on the likelihood of firm-branch matching: physical proximity is likely to alleviate informational frictions affecting banks' screening and monitoring costs (see Agarwal and Hauswald, 2010) as being close to clients eases the acquisition and the use of private information in informationally opaque credit markets.

Coefficients β_1 and β_2 control for heterogeneity in size and dynamics of bank branches documented in section 2.1. $\beta_1 > 0$ and $\beta_2 > 0$ would suggest that large and fast-growing branches are likely to be more efficient or visible and, conditional on distance, to offer better contract terms. Finally, coefficients β_3 and β_4 control for the direct impact of branch specialization (industry or size category) on firm's choice of banking partner: this relates to Paravisini

Table 1.1: Endogeneous Firm-Branch Matching

	(1)	(2)	(3)	(4)	(5)
Log Branch-Firm Distance (km)	-0.311***		-0.298***	-0.300***	-0.317***
	(0.001)		(0.001)	(0.001)	(0.001)
Branch size (t-1) × Firm size		0.126***	0.083***	0.081***	0.137***
		(0.001)	(0.001)	(0.001)	(0.001)
Branch growth (t-1)			0.108***	0.197***	0.233***
			(0.004)	(0.005)	(0.005)
Branch Industry spec. (t-1)				0.312***	
				(0.009)	
Branch Size spec. (t-1)					0.875***
					(0.009)
Observations	9,645,373	9,645,373	9,645,373	9,645,373	9,645,373
R-square	0.15	0.04	0.17	0.17	0.19

Notes: Probit estimation of different specifications of Equation (1.3). Dependent variable is an indicator function that takes the value 1 if firm i is borrowing from branch j at t. Spec. $\text{Sector}_{j,t-1}$ (respectively Spec. $\text{Size}_{j,t-1}$) equals 1 if branch j is specialized in lending to firms from the same sector (respectively same size) as firm i, the year before firm i entry. $\text{Size}_{j,t-1}$ is the log size of branch j, measured as total credit exposure and, alternatively, as the number of clients and Size $\text{Growth}_{j,t-1}$ is the growth rate of branch j, both measured the year before firm i entry. $\text{Distance}_{i,j}$ is simply the log distance between branch j and firm i. All explanatory variables are lagged. Standard errors clustered at the urban unit level.

et al. (2015), documenting that banks specialize in one export market and that specialization affects a firm's choice of new lenders and how to finance exports. $\beta_3 > 0$ (respectively, $\beta_4 > 0$) would mean that a given firm i is more likely to match with a bank branch j that has developed ex-ante a specific advantage in lending to firm' i industry (alternatively, the firm' i size category).

Results. Table 1.1 shows the marginal effects from the Probit estimation of different specifications of equation (1.3). Standard errors are clustered at the urban unit level. In every specification, all coefficients are statistically significant (at the 1 percent confidence level), and of the expected signs. Both physical proximity (α), branch size (β_1) and branch growth (β_2) increase the likelihood that a firm match with a bank branch. More interestingly, the actual existing portfolio of a branch shapes its future matches: a branch that specializes in lending to industry a is more likely to be chosen by a firm from industry a when it enters the credit market. To conclude, I find reduced-form evidence that bank branch heterogeneity does matter for small firms financing decisions; SMEs are more likely to search for and then match with branches that ex-ante exhibit a high-level of complementarity, which could

suggest the presence of search costs if firms devote time and ressources to gather information about potential banking partners before applying for a loan.

2.3. *The Geography of Bank Credit*

I describe the geography of credit flows and document new facts consistent with the presence of search frictions. In a frictionless world, meeting with many bankers is not costly and firms always benefit from borrowing from neighbouring banks [8], if possible, and inter-urban unit lending should be marginal. In particular, if a branch and a firm located in the same urban unit are simultaneously offering and applying for credit, a match is likely to occur locally. Therefore, we shoudn't observe a branch and a firm, while being located in the same urban unit, simultaneously lending and borrowing from another distant urban unit.

Simultaneous inter-submarket lending. In order to verify this assumption, I construct credit flows between each pair of urban units, using data from the French Credit Register between 1998 and 2005. Credit flows from i to j are defined as the total credit granted by all bank branches located in i to remote firms located in j: urban units may be lenders, borrowers or both. Table 1.10 presents the results.

[FIGURE 1.10 ABOUT HERE]

The first bar (in blue) indicates that more than 70 % of urban units *borrow* and *lend* simultaneously to remote submarkets. One could think of it as being the result of regional specialization. A firm located in a urban unit where branches are not able to offer a very specific type of credit, or to lend to a specific industry / firm size, may be forced to search for creditors elsewhere. Yet, the other blue bars indicates that a vast majority of urban units *borrow* and *lend* simultaneously the same type of credit to remote submarket: SMEs in urban unit a borrow long-term credit (LT) to branches in urban unit b, while branches in urban unit a simultaneously lend LT credit to SMEs in urban unit c.

Simultaneous two-way lending. More interestingly, I show that *two-way lending* occurs. Red bars on figure 1.10 indicate that more than 60 % of urban units simultaneously *borrow*

[8] A large body of work has highlighted the importance of geographical distance on firm-bank relationship, especially for small firms (Petersen and Rajan (1995), Hauswald and Marquez (2006)). Degryse and Ongena (2005) show that lenders located in the vicinity of small firms face significantly lower transportation and monitoring costs: *Banks derive market power ex ante from their relative physical proximity to the borrowing firms or ex post from private information they obtain about firms during the course of the lending relationship.* Consequently, most of the lending activity should concentrate within a city or an urban unit, between a firm and a branch in close proximity to each other.

and *lend* to the same distant submarket. Some bank branches located in Paris finance firms in Bordeaux, while some bank branches in Bordeaux finance firms in Paris. This fact is robust when I control for the type of credit and the type of firm (red bars 2 to 6).

2.4. Price Dispersion

Finally, inspired by the labor market literature, I document the fact that the law of one price does not hold in French credit submarkets. Using rich quarterly micro data on new loans to SMEs from the Sirius/M-Contran database, I show that credit rates exhibit a substantial dispersion within a time-bank branch-industry-department quadruplet, consistent with recent evidence on mortgage and consumer credit markets (see Argyle et al., 2019; Allen et al., 2019). Note that controlling for loans and borrowers characteristics does not affect my result. Formally, I estimate equation (1.4) that aims to explain the observed variation in loan prices:

$$\begin{aligned}\text{Interest rate}_{ijtu} = \text{Loan}_{ijtu} \cdot \rho_1 + \text{Firm}_{ijtu} \cdot \rho_2 \\ + \text{FE}_{s(u)} + \text{FE}_j + \text{FE}_t + \text{FE}_u + \epsilon_{ijtu}\end{aligned} \tag{1.4}$$

where i stands for the borrower, j for the bank branch, t for the quarter and u indicates the urban area in which both the firm and the bank branch operate. Loan_{ijtu} is a vector of loan characteristics (term in months, amount, type of rate: fixed or variable), Firm_{ijtu} is a vector of firm i characteristics (age, size, debt, investment grade, turnover). I sequentially add a bank (alternatively, a branch) fixed-effect, county fixed effect (i.e. French "Departments"), a sector (NACE Rev. 2 French classification) fixed-effect and a quarter fixed-effect. Table 1.2 shows the R^2 of an OLS regression for different specifications of equation (1.4) and for three categories of credit: equipment loans, credit lines and leasing.

As I am mostly interested in the explanatory power of these different groups of observable variables and fixed-effects, I only report the R^2 of each regression. The results indicate that, at best, the model accounts for 70% of the observed variance in credit prices, letting more than 30% (40% for leasing) of the variance unexplained, even when the model is saturated. Similarly, Cerqueiro et al. (2011) find substantial dispersion in loan rates for seemingly identical borrowers, using confidential Belgian data. The authors attribute this dispersion to information imperfections and asymmetries affecting credit markets and, among them, search costs. Note that if my reported R^2 is somehow larger than Cerqueiro et al. (2011) findings, this is due to the very strict inclusion of fixed-effects, notably bank branch FE, in my empirical model.

Table 1.2: Explaining Price Dispersion

R^2	(1) Time + Bank FE	(2) (1) + Dep. FE	(3) (2) + Sector FE	(4) (3) + Branch FE	(5) (4) + Loan	(6) (5) + Firm
Equipment Loans	0.628	0.647	0.652	0.672	0.689	0.699
Credit Lines	0.559	0.574	0.579	0.606	0.651	0.657
Leasing	0.491	0.511	0.533	0.533	0.538	0.553

Notes: R^2 from OLS estimations of equation (1.4). Dependant variable is the bank interest rate. Loan_{ijtu} is a vector of loan characteristics (term in months, amount, type of rate: fixed or variable), Firm_{ijtu} is a vector of firm characteristics (age, size, debt, investment grade, turnover). Column (1) include time and bank fixed effects. In columns (2) to (6), I sequentially add a county fixed effect (i.e. French "Départments"), a sector (NACE Rev. 2 French classification) fixed-effect and a quarter fixed-effect.

2.5. Survey Evidence

In this section, I document a series of evidence of search frictions from recent surveys that are consistent with empirical findings presented above. First, industry reports show that locating the right banking partner is not straightforward and that firms commonly multiply loan applications: the FED (2014) survey indicates that 3 applications are submitted on average (2.7 institutions contacted). Second, surveys highlight the fact that the application process for a loan is time-consuming: 33 hours are spent applying for credit on average according to the FED (2014) report, consistent to Infosys (2018) survey: *SMEs spend over 25 hours – simply on their loan request paperwork – and have to approach numerous banks with their application.* Among the SMEs, 26% (29%) deplore a difficult application process with large (small) banks.

Additionally, surveys highlight important transaction delays. Industry reports show delays in the range of 45-90 days between the application and the closing date. The Infosys (2018) survey indicates that 24% (29%) of SMEs report a long wait for the credit decision or funding with large banks (small banks), namely *high underwriting, transaction and search costs*. The OECD (2018) report "Enhancing SME access to diversified financing instruments" corroborates this conclusion: transaction costs are particularly high in relative terms for micro-entreprises, start-ups, young SMEs. This costly search, among other factors, may result in firm resignation: 37% of businesses appear to give up their search for finance and cancel their spending plans after their first rejection (BIS/BMG Research, 2018).

3. Model

Motivated by empirical evidence I document on search frictions in French credit markets, I present a partial equilibrium model of firm-bank matching and inter-regional credit flows based on Eaton and Kortum (2002), that incorporates realistic geographic aspects. The model features two-sided heterogeneity – bank branches and firms – and information frictions of two kinds; First, informational asymmetries affect banks ability to screen and monitor projects. Second, search frictions hinder firms ability to locate and match with the right financing partner, as in Eaton et al. (2018) and Lenoir et al. (2018).[9] The model captures the key empirical evidence presented in section 2. In the following sections, I start by summarizing the main assumptions; Then, I derive analytical predictions on aggregate credit flows and firm-bank matching.

3.1. *Setup*

There are a large number of local submarkets in the economy, indexed by u = 1,...,N , each inhabited by an exogenous mass of entrepreneurs (SMEs) and bank branches. In what follows, I use u to refer to the submarket in which a bank branch is located (the *origin* submarket) and v to refer to the submarket in which the branch customer is located (the *destination* submarket) . In this economy, a single *good* is consumed by entrepreneurs and provided by bank branches into perfectly substitute varieties: bank credit.

Supply side. There is a continuum of bank branches in each submarket u, of measure $N_u = S_u \cdot z_{min}^{-\theta}$, with S_u indicates the size of the submarket. Bank branches produce and provide credit with a single factor constant return-to-scale production function[10]. For the sake of simplicity, I make no distinction between short-term and long-term loans and I do not model more complex credit types as leasing or factoring. Bank branches operating in submarket u incur an exogenous input unit cost c_u, that encompasses the branch office rent, loan officer wage or marketing expenses, among others. The productivity of a bank branch b_u located in submarket u is independently drawn from a Pareto distribution of parameter

[9] This model is inspired by the recent trade literature that emphasise the role of search frictions in international goods market. Here, I model bank credit as a special kind of *good* that require buyer-supplier search and matching and involve no traditional transportation cost.

[10] Constant return-to-scale technology for bank branches seems to be a reasonable assumption as it has been documented for large financial institutions (McAllister and McManus, 1993); However, increasing return-to-scale appears to better fit the data for most US banks, a fact further documented in Wheelock and Wilson (2012). Using increasing rather than constant return-to-scale does not affect the predictions of the model.

θ and support $[z_{min}, +\infty[$:

$$z_{b_u} \sim \text{Pareto}\ (z_{min}, \theta) \tag{1.5}$$

Note that this Pareto assumption is data driven (see section 2.1): branch size – measured as its total credit exposure – is used as a proxy of branch productivity, which I can't measure accurately. Thus, the number of bank branches located in submarket u that can provide credit with efficiency above z writes $N_u(z) = S_u \cdot z^{-\theta}$.

Bank branches located in u additionally incur a variable cost d_{uv} when lending to a firm located in a remote submarket v, which is a function of the physical distance between submarkets u and v. This *iceberg cost* encompasses the fact that branches located closer to borrowing firms enjoy significantly lower i) transportation, ii) screening and iii) monitoring costs (Degryse and Ongena, 2005). In line with theoretical location models (Hotelling, 1929; Salop, 1979) that include distance-related transportation costs, transportation and monitoring costs increase with borrower-lender distance because of additional communication costs and time-consuming transport for the loan officer when visiting its remote clients. Screening costs, stemming directly from asymmetric information (Hauswald and Marquez, 2003), also decrease with firm-branch distance, such as the precision of the signal about borrower's quality. Thus, bank branches market power arises from this proximity to local borrowers and erodes over distance.

Demand side. Each submarket u is populated with a continuum of ex-ante heterogeneous entrepreneurs (or SMEs) with an investment project I, of size normalized to one, and no cash. Entrepreneurs differ in their productivity z_e. The productivity of an entrepreneur e_u located in submarket u, is independently drawn from a Pareto distribution of parameter γ and support $[z_{min}, +\infty[$:

$$z_{e_u} \sim \text{Pareto}\ (z_{min}, \gamma) \tag{1.6}$$

such as the number of firms in a submarket writes $F_u = S_u \cdot z_{min}^{-\gamma}$[11]. In order to start their investment project, entrepreneurs need to raise external finance from banks – which is the only source of external finance available for small firms. Because of search frictions, it is difficult to locate the right banking partner; as a consequence, an entrepreneur has to undergo costly search process.

[11] I remain agnostic about the entrepreneur production function that can be either Cobb-Douglas or CES, without altering the model predictions.

Search and matching. I build on Eaton et al. (2018) where matching between buyers and sellers is random. Each entrepreneur meet with a discrete number of bank branches, some located in their own submarket, some located remotely. This random search process is a reduce form for the active search for banking partners: entrepreneurs need to gather information about bank branches characteristics, contact loan officers and, finally, physically meet with them to get a price quote.

Formally, the discrete number of branches met in submarket u is drawn into the distribution $N_u = S_u \cdot z_{min}^{-\theta}$. This implies that the number of branches met with efficiency higher than z is drawn in $N_u(z)$. As a consequence, the set of potential lenders drawn by entrepreneur e_u is the random variable Θ_{e_u}, which is the sum of potential banking partners met in each of the N submarkets. Θ_{e_u} reflects the strength of search frictions affecting the submarket u; in a frictionless world, each entrepreneur from u would meet with all bank branches in the economy and, in turn, would end up applying for a credit from their *optimal* banking partner (i.e. the first-best match)[12]. In decentralized credit markets with search frictions hindering the number of meetings and price quotes, the first-best match is not always feasible as an entrepreneur may never meet with the right loan officer.

In Eaton et al. (2018) and Lenoir et al. (2018), there is no firm heterogeneity and the likelihood to meet with a supplier from v is the same for all the firms in u. I instead assume that entrepreneurs heterogeneity matters and reflects the fact that more productive entrepreneurs incur lower search costs. I model the search process as independent draws in the distribution of bank branches; each bank branch b_u located in u has the probability $z_{e_v}\kappa_{uv}$ to be drawn by entrepreneur e_v located in v. z_{e_v} stands for the firm productivity (normalized such that $z_{e_v} \in [0, 1]$). κ_{uv} (also $\in [0, 1]$) can be interpreted as a pair-specific u-v inverse measure of the strength of search frictions. Formally, $\mathbb{P}[b_u \in \Theta_{e_v}] = z_{e_v}\kappa_{uv}$ and $\Theta_{e_v}(u)$, the number of bank branches from u met by an entrepreneur from v, follows a binomial law such that:

$$\mathrm{Card}(\Theta_{e_v}(u)) = z_{e_v}\kappa_{uv} \cdot S_u z_{min}^{-\theta} \tag{1.7}$$

Under the Poisson limit theorem, the binomial law of parameters $(z_{e_v}\kappa_{uv},\ S_v z_{min}^{-\theta})$ can be approximated by a Poisson law of parameter $z_{e_v}\kappa_{uv} \cdot S_v z_{min}^{-\theta}$; this approximation is used in

[12]Note that, without search frictions, all entrepreneurs located in the same submarket may end up applying for credit from the exact same bank branch, if only branch productivity matters (Cerqueiro et al. (2011) shows that branch heterogeneity goes far beyond productivity). This will lead to positive assortative matching and directed search.

the rest of the analysis. This modelling has two major implications. First, more productive firms will, mechanically, meet with more bank branches, not only locally but also in distant submarkets. De facto, productive entrepreneurs are more likely to find a *good* match among Θ_{e_v} while entrepreneurs with a low productivity may end up with only a few bad quotes. Second, heterogeneity in κ_{uv} across submarkets implies that entrepreneurs' search will be biased geographically toward submarkets in which search frictions are low. An important feature of the search process is that bank branches heterogeneity does not affect the probability of meeting; in particular, there is no directed search toward the most productive branches. I argue that this is a reasonable assumption, given the fact that branch characteristics (specialization, growth rate, etc.) as well as loan officer background, preferences and bargaining ability are difficult to assess from an outsider. This echoes Cerqueiro et al. (2011) notion of loan officer *discretion* in the loan rate setting process, especially for small and opaque businesses. Gathering information about How much a bank branch will be a good fit turns out to be costly and complex, so I assume that branch characteristics are ex-ante unobserved and do not affect the meeting probability.

Conditional on meeting with a loan officer in u, entrepreneur e_v pitches its investment project and get a price quote. I assume a simplistic bank pricing strategy (bank branches always price at their marginal cost, as in a perfect competition framework) and a reduced-form cost function that depends on both branch and entrepreneur productivity, exogenous unit cost c_u and transportation cost d_{uv}. The interest rate offered by bank branch b_u to lend to the entrepreneur e_v writes $r_{b_u,e_v} = \frac{c_u d_{uv}}{z_{e_v} z_{b_u}}$[13]. The interest rate decreases with branch-firm distance and the unit cost of production, and is negatively correlated to both branch and entrepreneur productivity. It summarizes and aggregates all the costs linked to information asymmetries faced by the loan officer: screening, monitoring and transportation costs discussed above.

The assumption of bank branches pricing at their marginal cost is strong, in particular in a context of credit markets subject to informational asymmetries. As pointed out in Lenoir et al. (2018), an alternative is to assume that bank branches compete à la Bertrand: the branch that offers the best contract terms doesn't price at its marginal cost, but equals the marginal cost of the branch with the *second best* offer. Another alternative is to assume a Nash bargaining equilibrium in which the branch with the best offer and the entrepreneur share the surplus of the match. Under the assumption of inelastic demand, competition

[13]The bank branch b_u program writes: $\Pi_u = z_{match} K r - K c_u d_{uv}$ with $z_{match} = f(z_{b_u}, z_{e_v})$. I assume a multiplicative reduced-form expression for $f(z_{b_u}, z_{e_v})$ which is quite common in the labor literature where the productivity of a worker-firm match is simply the product of both productivities.

à la Bertrand and Nash bargaining do not affect the model predictions about firm-branch matching. Since firm-branch matching is the main outcome of my model, I keep the marginal cost pricing assumption for the sake of simplicity.

After meeting with all the loan officers in Θ_{e_v}, the entrepreneur e_v decides to match with the one offering the best contract terms (i.e. lowest interest rate). The interest rate paid by e_v writes:

$$r_{e_v} = \underset{b_u \in \Theta_{e_v}}{\text{argmin}} \{ \frac{c_u d_{uv}}{z_{b_u} z_{e_v}} \} \tag{1.8}$$

The poisson search process combined with the Pareto distribution of brank branches size allows an analytical formula for r_{e_v}. Eaton et al. (2018) demonstrates that the assumption of Poissons draws into a Pareto distribution delivers a Weibull distribution for the minimum interest rate introduced in equation 1.8. Formally:

$$\mathbb{P}\big(r_{e_v} \leq r\big) = W_{e_v}(r) = 1 - \exp\Bigg(- r^{\theta} z_{e_v}^{\theta+1} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv} \Bigg) \tag{1.9}$$

Conditional on r fixed, entrepreneurs in submarket v obtain on average a best offer if the level of competition is high (i.e. $\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}$ is large), due to the proximity with vast and crowded submarkets. In the same vein, the lower the search frictions faced by entrepreneurs in v (i.e. the greater κ_{uv}) the better the contract terms will be, on average. Finally, the entrepreneur's productivity directly impacts on the likelihood to be matched with a branch that offers a low interest rate because more productive entrepreneurs draw a larger set Θ_{e_v} of potential lenders that somehow compensates the adverse effect of search frictions. Note that a larger θ also alleviate the negative impact of search frictions, but mostly in favor of less productive entrepreneurs: indeed, less heterogeneity between bank branches advantages the entrepreneurs with the least price quotes.

3.2. Predictions

In this section, I derive a number of theoretical predictions about i) the magnitude of aggregate credit flows between any two submarkets, ii) firm-branch matching and iii) the number of clients by bank branch along the distribution of branch's productivity. I then investigate how a shock on search frictions modifies those predictions.

Aggregate Credit Flows

In section 2.1 and 2.3, I document the existence of inter submarket lending: while most of the credit is distributed locally, with borrowers and lenders located in the same urban unit, some branches operate in remote submarkets. Let Π_{uv} be the share of credit granted in submarket v by bank branches located in submarket u (over the total credit borrowed by firms in v). As all investment projects share the same size, normalized to one, the expected share of credit distributed in v by branches located in u is the sum over all entrepreneurs in v of a dummy variable equal to one if the entrepreneur is matched with a branch in u, zero otherwise:

$$\Pi_{uv} = \frac{\sum_{e_v=1}^{F_v} \mathbb{I}\{\mathrm{M}(e_v) = u\}}{\sum_{e_v=1}^{F_v} 1} \tag{1.10}$$

where $\mathrm{M}(e_v) = u$ indicates that entrepreneur e_v decides to match with a branch from u. Lenoir et al. (2018) shows that using the law of large numbers, Π_{uv} is equal to the expected value of $\mathbb{I}\{\mathrm{M}(e_v) = u\}$ across entrepreneurs in v, which is the probability that the best contract terms offered to any entrepreneur in v comes from a branch u. Here, a crucial condition is that random variables $\mathbb{I}\{\mathrm{M}(e_v) = u\}$ are independent and identically distributed, which is straightforward if entrepreneurs are ex-ante identical. In my case, with entrepreneur's heterogeneity, I show that this condition holds as the likelihood to ultimately match with a submarket u does not depend on z_{e_v}. Thus, equation 1.10 writes:

$$\Pi_{uv} = \mathbb{E}_{e_v}\Big[\mathbb{I}\{\mathrm{M}(e_v) = u\}\Big] = \mathbb{P}\Big[\{\mathrm{M}(e_v) = u\}\Big] \tag{1.11}$$

I consider a level of price r (and a level of productivity z_{e_v}) fixed. Let $D_{u,e_v}(r)$, be the number of branches from u met by entrepreneur e_v that propose an interest rate below r, formally $D_{u,e_v}(r) = z_{e_v}^{\theta+1} r^{\theta} S_u (c_u d_{uv})^{\theta} \kappa_{e_v}$. Then:

$$\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \frac{\int_{D_{e_v}(r)>0} \mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r, D_{e_v}(r)\Big]\, \mathrm{dF}(D_{e_v}(r))}{\int_{D_{e_v}(r)>0} 1\, \mathrm{dF}(D_{e_v}(r))} \tag{1.12}$$

In equation 1.12, the numerator can be interpreted as the discrete sum over all the total possible number of draws $D_{e_v}(r) = n$, of all the possible combinations of draws from u, i.e.

n_u, among n. This leads to:

$$\begin{aligned}
\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \\
&\sum_{n=1}^{+\infty}\sum_{n_u=0}^{n}\Bigg[\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r, D_{u,e_v}(r) = n_u, D_{k\neq u,e_v}(r) = n - n_u\Big] \\
&\times \mathbb{P}\Big[D_{u,e_v}(r) = n_u\Big] \times \mathbb{P}\Big[D_{k\neq u,e_v}(r) = n - n_u\Big]\Bigg] \\
&\times \mathbb{P}\Big[D_{e_v}(r) > 0\Big]^{-1}
\end{aligned} \tag{1.13}$$

After some calculations, I obtain the following equation for Π_{uv}, at r fixed:

$$\begin{aligned}
\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = &\frac{z_{e_v}^{\theta+1}\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{z_{e_v}^{\theta+1}\sum_{k=1}^{N}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}} \\
&\times 1 - \exp\left(-r^\theta z_{e_v}^{\theta+1}\sum_{u=1}^{N} S_u\cdot(c_u d_{uv})^{-\theta}\kappa_{uv}\right) \\
&\times \mathbb{P}\Big[D_{e_v}(r) > 0\Big]^{-1}
\end{aligned} \tag{1.14}$$

Note that $\mathbb{P}[D_{e_v}(r) > 0] = \mathbb{P}[r_{e_v} < r]$, the probability for the minimum price quote to be lower than p, for which an analytical formula is given by equation 1.9, leading to the following simplification for equation 1.13 :

$$\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \frac{z_{e_v}^{\theta+1}\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} z_{e_v}^{\theta+1}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}} \tag{1.15}$$

First, equation 1.15 indicates that the likelihood for an entrepreneur located in v to match with a bank branch from u does not vary along the distribution of productivity. When z_{e_v} increases, the number of branches drawn by e_v in u mechanically increases but so do the number of branches drawn in others competing submarkets ($\forall k \neq u$), which results in a constant probability of matching with u. Entrepreneur productivity only impacts the contract terms, not the destination of the match. Second, under the assumption that Pareto distributions of bank branches productivity share the same shape parameter θ across submarkets, $\mathbb{P}[\{\ \mathrm{M}(e_v) = u\}|r]$ is the same for each price quote r. Thus, the structural expression for the share of credit distributed in v by branches located in u is:

$$\Pi_{uv} = \mathbb{E}_{ev}\Big[\{\mathbb{I}\ \mathrm{M}(e_v) = u\}\Big] = \frac{\kappa_{uv} S_u (c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}} \tag{1.16}$$

Two forces are at stake: i) the relative magnitude of search frictions between submarkets u and v with respect to the magnitude of search frictions affecting all the other potential submarkets $k \neq u$ and ii) the relative size and efficiency of the submarket u compared to submarkets sizes and efficiencies in the rest of the economy. From, equation 1.16, I derive two predictions about aggregate credit flows and the impact of a shock on search frictions κ_{uv}.

Prediction 1: Gravity Equation for Bank Credit. As shown in Lenoir et al. (2018), a log-linearization of equation 1.16 delivers a gravity equation for the share of credit distributed in v by branches located in u:

$$\log \Pi_{uv} = \log \kappa_{uv} - \theta \cdot \log d_{uv} + \mathrm{FE}_v + \mathrm{FE}_u \tag{1.17}$$

where FE_u stands for log $S_u c_u^{-\theta}$ and FE_v equals $-$log $\sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}$. Gravity equations are not common in the finance literature, with notable exceptions for cross-border equity flows (Portes and Rey, 2005), bonds ans bank holdings (Coeurdacier and Martin, 2009). Recently, Okawa and van Wincoop (2012) proposed a theoretical foundation of a gravity equation for cross-border asset holdings gravity including financial frictions in the form of informational asymmetries about assets future returns. To the best of my knowledge, this paper is the first to propose a structural gravity equation for within-country bank credit flows with two types of informational asymmetries.

Prediction 2: A Shock on Search Frictions. I investigate the effect of a reduction of bilateral search frictions – e.g. the development of Broadband Internet and online banking services – on aggregate credit flows. First-order condition of equation 1.17 with respect to κ_{uv} leads to:

$$\frac{\partial \ln \Pi_{uv}}{\partial\, \kappa_{uv}} = \underbrace{\frac{\partial \ln \kappa_{uv}}{\partial\, \kappa_{uv}}}_{(a)} + \Big[\underbrace{-\frac{\partial \ln\, \sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}}{\partial\, \kappa_{uv}}}_{(b)} \Big] \tag{1.18}$$

Two opposite mechanisms are at stake. First, a reduction of search frictions has a direct and strictly positive effect (a) on bilateral credit flows. This *connectivity effect* reflects the fact that it becomes less costly for entrepreneurs located in v to gather information about bank

branches and loan officers in u and to meet with them. Formally, the likelihood of meeting with a bank branch from u increases for each entrepreneur in v; i.e. there will be in average more bank branches from u in Θ_{e_v}. Second, the *competition effect* (b) captures the increasing competition between bank branches, conditional on being met, induced by the higher number of potential banking partners. The expression (b) rewrites $\frac{S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^N \kappa_{kv} S_k(c_k d_{kv})^{-\theta}} \geq 0$, such that the *competition effect* is negative and may compensate the direct effect of *connectivity*. Equation 1.18 simplifies in:

$$\frac{\partial \ln \Pi_{uv}}{\partial \, \kappa_{uv}} = \frac{1}{\kappa_{uv}} - \frac{S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^N \kappa_{kv} S_k(c_k d_{kv})^{-\theta}} \geq 0 \tag{1.19}$$

Note that the effect of a reduction of search frictions κ_{uv} in heterogeneous across submarkets u. The larger or closer a submarket, the smaller the total impact. This captures the fact that large and nearby submarket already benefits from a *visibility* advantage; entrepreneurs in v easily meet with bankers from those very accessible and visible submarkets which, in turn, are able to offer attractive contract terms. In contrast, small and remote submarkets benefit more for a reduction of search frictions.

Branch-Entrepreneur Matching

I investigate the matching process between an entrepreneur located in v and any bank branch located in the N submarkets. I derive predictions about (i) the number and the quality of entrepreneurs that ultimately match with a particular bank branch and (ii) the impact of a reduction of search frictions on the matching equilibrium. Both predictions can be confronted to firm-branch relationship data from the Credit Register and to motivating empirical work presented in 2.

Prediction 3: Positive Assortative Matching. I consider a bank branch located in u and its likelihood $F_{b_u}(e_v)$ to lend to the entrepreneur e_v located in v as a result of the search and matching process. $F_{b_u}(e_v)$ can be decomposed as the likelihood for entrepreneur e_v to draw and meet with b_u and then, the likelihood for b_u to be the lowest cost supplier. Formally:

$$\begin{aligned} F_{b_u}(e_v) &= \mathbb{P}\Big(b_u \in \Theta_{e_v}\Big) \times \mathbb{P}\Big(\underset{\Theta_{e_v}}{\text{argmin}}\{\frac{c_u d_{uv}}{z_{b_u} z_{e_v}}\} = b_u\Big) \\ &= \mathbb{P}\Big(b_u \in \Theta_{e_v}\Big) \times \Big(1 - \mathbb{P}(r_{e_v} < \, r_{bu})\Big) \end{aligned} \tag{1.20}$$

From equation 1.9 for minimum price distribution, I have an analytical formula for $\mathbb{P}(r_{e_v} \leq$

r_{bu})[14]. By definition of the random search process, $\mathbb{P}(b_u \in \Theta_{e_v}) = z_{e_v}\kappa_{uv}$. Thus, the model delivers the following expression for the likelihood that e_v ultimately decide to borrow from b_u, namely $F_{b_u}(e_v)$:

$$F_{b_u}(e_v) = z_{e_v}\kappa_{uv} \times \exp\left(- (c_u d_{uv})^{\theta} z_{b_u}^{-\theta} z_{e_v} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv} \right) \tag{1.21}$$

The likelihood of a match between branch b_u and entrepreneur e_v strictly increases in z_{bu} and z_{e_v}. The branch productivity has a simple and direct effect *via* the attractiveness of the price quote while the entrepreneur productivity has two distinct impacts: a direct positive effect on the meeting likelihood that appears in the first part of equation 1.21 and a negative but smaller effect – in the exponential term – that captures a competition effect (conditional on meeting, the likelihood of being the lowest cost bank branch decreases with the number of other branches met). This indicates a positive assortative matching between very productive branches and entrepreneurs. On the contrary, low-productivity bank offices are likely to match only with unproductive entrepreneurs that do not enjoy a large set of potential partners. This is consistent with empirical evidence from the Probit regression presented in section 2: the product of branch and firm size is positively correlated with branch-firm matching, as the branch growth rate is.

Prediction 4: Firm-Branch Matching and Search Frictions. The first-order condition of equation 1.21 with respect to κ_{uv} indicates How the matching process is affected by a shock on bilateral search frictions between any two submarkets u and v.

$$\begin{aligned}\frac{\partial \ln F_{b_u}(e_v)}{\partial\, \kappa_{uv}} &= \underbrace{\frac{\partial \ln \left(z_{e_v}\kappa_{uv}\right)}{\partial\, \kappa_{uv}}}_{(a)} \\ &\quad \underbrace{- (c_u d_{uv})^{\theta} z_{b_u}^{-\theta} z_{e_v} \sum_{k=1}^{N} S_k \cdot (c_k d_{uk})^{-\theta} \frac{\partial\, \kappa_{kv}}{\partial\, \kappa_{uv}}}_{(b)}\end{aligned}$$

Similar to *Prediction 2*, the impact of a reduction of search frictions is twofold, with a direct and positive *connectivity effect* (a) and an indirect and negative *competition effect* (b). The *connectivity effect* captures the enhanced visibility of the bank branch. The *competition*

[14] I made the usual approximation that $\mathbb{P}(r_{e_v} \leq r_{bu}) \approx \mathbb{P}(r_{e_v} < r_{bu})$

effect reflects the fact that, conditional on being drawn, it becomes more difficult to offer the lowest price quote. Equation 1.22 offers a reduced-form expression that summarizes both effects:

$$\frac{\partial \ln F_{b_u}(e_v)}{\partial \kappa_{uv}} = \frac{1}{\kappa_{uv}} - \frac{S_u z_{e_v}}{z_{b_u}^{\theta}} \tag{1.22}$$

The impact of search frictions varies along the distribution of branch productivities. High-productivity branches (high z_{b_u}) benefit more form the reduction of search costs as the direct impact dominates the competition effect: those large and efficient branches located in v now meet with much more entrepreneurs from v and, conditional on being met, enjoy a dominant position compared to smaller and less productive branches when it comes to offer attractive contact terms. Their likelihood to be the lowest price quote and, then, to ultimately be chosen, increases. On the contrary, low-productivity branches that previously benefited from the low level of competition (i.e. few other branches drawn from u) are now exposed to a tougher rivalry.

4. Empirical Context

In this section, I provide a description of my empirical strategy. I exploit the staggered deployment of Broadband Internet in France between 1998 and 2005 as a large scale quasi-natural experiment to study the impact of a reduction of search frictions on credit markets. First, I present the technological and institutional context of Broadband Internet diffusion in France. Second, I document the adoption of ICTs by French banks and How high-speed Internet is likely to reduce search frictions for SMEs. Third, I discuss the identification strategy and propose a new IV for Broadband Internet deployment.

4.1. The Rise of Online Banking

In the early 2000's, the large diffusion of ICTs represented a profound change for the banking industry and Broadband Internet was the catalyst for this numerical transformation. As digitization proceeded apace, the dimensions of banks' digital evolution varied among segments and products. Digitization became the norm for retail credit processes with personal-loan applications submitted with a few swipes on a mobile phone, and time to cash can now be as short as a few minutes (McKinsey, 2018). Not only transaction costs decreased: the rise of online price comparison services and brokers allows individuals to search for the best

banking partner in a faster and more efficient way[15]. The combination of both transaction and search costs reduction resulted in a severe disruption of search frictions for individuals.

Regarding corporate credit and SME lending, the situation is mixed. The loan officer remains the most relevant interlocutor and the ultimate decision-maker in SME lending, as He is ideally suited to understand client's specific needs and characteristics as well as local market and industry performance, leaving little room for automation. The complexity of loan pricing for SMEs also prevents the use of online brokers or interest rate comparison websites. However, ICTs affected many aspects of the firm-bank relationship, especially for SMEs: recent survey on UK SMEs (Ernst and Young, 2018) shows that financial services used by SMEs are mainly online banking (85%) and second, branch-based banking (43%), emphasizing the growing importance of digitization. Among other examples, emails allow firms to easily contact a loan officer, online data and document sharing speeds up the meeting process and reduces transaction costs. Bank websites are showcases designed to attract new customers and provide information about financial products. Finally, customer online areas facilitate communication.

In order to document the fact that French banks had started their digital transition process at the beginning of the 2000's, I gather new data on large French bank adoption of ICTs; First, I check the existence of the bank website with a firm customer area before 2000 (using the waybackmachine.com website). Second, I collect the exact date of the domain name creation (available on nom-domaine.fr). The six largest French banks, which represent around 90% of the total amount of credit granted to firms, were already active online at the beginning of the 2000's, with a sophisticated website, while their domain name were registered in average in the mid-90's, showing an early preoccupation for online visibility. In particular, the average website demonstrate the willingness of banks to improve the accessibility of basic but crucial information to their future SME clients: in a few clicks, it was possible for a new client to get an appointment, find all the bank branches in the area or to ask for information about financial products and services. For each bank branch, the phone number, the contact email and the physical address were immediately available. This represent a sharp reduction of search costs born by the entrepreneur.

[15]Mortgage lending is more complex due to regulatory constraints, yet banks in many developed markets have managed to digitize large parts of the mortgage journey. More than one bank has set an aspiration to automate 95 % of retail underwriting decisions.

4.2. The diffusion of Broadband Internet in France

The ADSL technology

The ADSL (Asymmetric Digital Subscriber Line) is a data communication technology that enables fast data transmission over copper telephone lines (much faster than what a conventional voiceband modem could provide). In the ADSL technology, bandwidth and bit rate are said to be asymmetric, meaning greater towards the customer premises (downstream) than the reverse (upstream). Eligibility for ADSL depends on the distance between the final customer (e.g. home or office) and the Local Exchange (LEs), since the intensity and the quality of the analog signal decreases as it is routed over the copper lines. Local Exchanges are the telephone exchanges owned by the incumbent operator France Télécom (later renamed Orange) into which subscribers' telephone lines end. As of 2008, there were about 17 000 LEs spread throughout the country. Initially dedicated to the telephone network, LEs are essential for Internet users who subscribe to ADSL as they aggregate local traffic and then direct it via the so-called backbone (i.e. higher levels of the network) towards the world wide web. A key feature of the ADSL technology is that one can supply high-speed Internet by upgrading the LE while relying on the existing (copper) local loop to connect the premises of the final customers. The upgrading involves the installation of an equipment inside the LE called a DSLAM (Digital subscriber line access multiplexer) that is required in order to recover the data transmitted via ADSL on the local copper loop and adapt it so it can be transmitted to the higher levels of the network (which are typically relying on optical fiber). The upgrading of local LEs is the key source of variation I will use in my empirical analysis.

The ADSL roll-out in France

The ADSL technology became popular during the 1990s, as many OECD countries were planning the expansion of services related to information and communications technology. In the early 2000s in France, the deployment of the technology beyond France's largest cities was slow. The causes for this staggered deployment are multiple. First, France Telecom (FT), the monopolistic telecom supplier at the time and still the main supplier today, was unsure as to whether it was going to be able to make the upgraded infrastructure available to new competitors with a positive markup or not. The uncertainty regarding the wholesale price FT was going to be able to charge made the firm reluctant to upgrade LEs beyond the largest cities (see Sénat, 2002, p.232). This uncertainty was lifted after a series of decisions by the regulatory agency set the conditions of that wholesale market (Arcep, 2002). Moreover,

at the same time France Telecom had to invest massively in upgrading its LEs to ADSL, it went through an debt crisis which ended with what was essentially a government bailout in 2002. One can find anecdotal evidence of the impatience of the French government in accounts of Parliamentary debate (at the Senate) regarding the excessively slow expansion of broadband internet (Tregouet, 2001) and the difficult cooperation between the French government (the Ministry in charge of the Industry) and France Télécom. Under the impulse of the government – which increased its stake in the firm during the 2002 bailout of the firm – France Telecom pledged in 2003 to cover 90% of the French (metropolitan) population by the end of 2005, i.e. all local exchanges (LEs) with more than 1000 lines, for a total investment of 600 M euros (750 M euros in 2018 prices) (Telecom, 2003).

Overall, the account of the broadband expansion in France over the period suggests that it was gradual due to uncertainty regarding the capacity of France Telecom to undergo the investment until 2002. After 2002, with the strong impulse of the government, France Telecom started covering more secondary areas with a focus on the overall number of lines per LE with only limited attention paid to local economic potential. While accelerated, the coverage remained gradual due to operational limits on the part of FT and took about 2 more years than anticipated in 2003.

4.3. Identification and Instrumental Variable

Similar to Magouyres et al. (2019), my identification strategy is based the gradual diffusion of the new technology in different LEs over space and time. Note that the question of what were the criteria for deciding to *treat* one LE before another has been studied in Magouyres et al. (2019). Empirical evidence shows that the main determinant of Broadband-Internet expansion was the city-level population density, with no role for levels or trends in the economic patterns of the city, and was slow down by the sunk cost of upgrading the infrastructure, consistent with statistical analysis[16].

In this paper, I propose an instrument variable strategy based on a theoretical optimal investment plan for infrastructure upgrading. The ADSL technology combines local copper loops and a large optic fiber network. Thus, when France Télécom decided to connect a specific city with Broadband Internet, the total cost of the project was twofold: the cost

[16]In particular, the authors highlight the fact that broadband expansion occurring to maximize population coverage with no special consideration for economic potential is strongly supported by a statistical analysis of the determinants of broadband coverage that is carried out in their paper. Different to their paper, my treatment is continuous and I can't rely on a dynamic event-study approach and check for the pre-trends. For this reason, I propose an IV strategy for the timing of Broadband expansion.

of upgrading the LE and the connection cost between the LE and the global optical cable network, which depends on the physical distance between the LE and the closest optic fiber cable. On the other hand, the *gain* for the internet supplier to upgrade a LE was proportional to the number of inhabitant newly connected. My optimal theoretical investment plan opposes the connection costs to the connection gains for each city. The gains are the number of potential clients (measured before 1999) reached consequent to a LE upgrading and I use the distance (in km) to the closest optic fiber as a proxy of the connection costs.

[FIGURE 1.11 ABOUT HERE]

A key feature of this instrumental variable strategy lies in the exogeneity of the distance between a LE and the closest optical fiber cable, as the optic fiber network construction was anterior to the ADSL expansion. The network has been built, in part, by other economic actors, before 1998 and for a completely different purpose. Indeed, highway firms and the French railway company (SNCF) installed optic fiber cables along the lines (respectively, the roadsides) for fast data transmission: surveillance videos, internal communications, etc. France Télécom leased the existing infrastructure to those company in order to faster Broadband expansion. Figure 1.11 displays the location of around 13,000 Local Exchanges, highways and railroads already existing before the beginning of Broadband Internet expansion in France. For each LE, I compute the shortest geodesic distance to a highway or a railroad, and use this as a proxy of connection costs.

[FIGURE 1.12 ABOUT HERE]

Formally, I predict the *optimal connection rank* $\hat{R}_i$ for the Local Exchange i, only taking into account two presumably exogenous measures of costs and gains. I use this optimal connection rank $\hat{R}_i$ in place of the observed rank R_i to predict the theoretical year of connection, and, thus, $\hat{Z}_{it}$. As a consequence, the optimal connection rank is not polluted by concomitant or correlated economic shocks that may affect the connection timing and only depends on preexisting and time-unvarying city characteristics:

$$R_i = \alpha + \beta_1 \cdot \text{Density}_{i,1998} + \beta_2 \cdot \text{Shortest Distance}_{i,1998} + \epsilon_i$$
$$\hat{R}_i = \hat{\alpha} + \hat{\beta}_1 \cdot \text{Density}_{i,1998} + \hat{\beta}_2 \cdot \text{Shortest Distance}_{i,1998}$$

Figure 1.12 shows the rank correlation between optimal versus observed connection ranks. The combination of exogenous connection *gains* and *costs* have a strong predictive power,

with a R-square close to 0.70. Finally, the mapping between the predicted connection rank $\hat{R}_i$ and the connectivity variable $\hat{Z}_{it}$ follows the correspondence between R_i and Z_{it} observed in the data, such that $\hat{Z}_{it}$ and Z_{it} also display a strong positive correlation. Finally, I define the degree of connection between two urban units as $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My instrument variable strategy delivers a similar measure of between urban unit connectivity named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$

5. Empirical Approach

In this section, I describe my empirical approach. I test the three main predictions from the model: (i) inter-regional credit flows follow a gravity equation, (ii) the Broadband Internet roll-out affects inter-regional flows through a reduction in search frictions and finally (iii) the Broadband Internet roll-out affects firm-branch matching.

5.1. Gravity Equation for Aggregate Credit Flows

I use data on bilateral credit exposure from the Credit Register, aggregated at the urban-unit level for all single-city SMEs in my sample, in order to compute bilateral credit shares Π_{uvt}. Π_{uvt} is the amount of credit granted by bank branches located in u to SMEs located in v, over the total credit stock of SMEs in v. Π_{uvt} lies in $[0, 1]$, 0 indicates that none of the firms located in v borrow credit from a bank branch located in u at date t (i.e. the aggregate credit flow from v to u is null). On the contrary, Π_{uvt} equals 1 implies that firms in v are fully financed by branches in u. In an economy without inter-regional exchanges, all Π_{uvt} would equal 0, except Π_{uut} as all firms would be financed by local bank branches. Alternatively, I use other measures of flows to distinguish between the extensive and the intensive margin: the total amount of credit granted to firms in v by branches in u, the average loan granted, the number of firms in v financed by u or the share of firms financed by u.

Baseline Specification. A very broad literature in international trade studies the gravity equation and its estimation (see Head and Mayer, 2014 for an overview). Santos Silva and Tenreyro (2006, 2011) show that the Pseudo-Poisson Maximum Likelihood (PPML) estimator, introduced by Gourieroux et al. (1984), is a promising workhorse for the estimation of gravity equations, in particular in the presence of many zeros. It is perfectly suited for the estimation of multiplicative models, without log-linearization of the dependent variable.[17]

[17] The PPML estimator identifies the coefficients using the same first-order conditions that are used by

I adopt this standard approach and rely on the new estimator for pseudo-poisson regression models with multiple high-dimensional fixed-effects developed by Correia et al. (2019). Formally, I estimate equation 1.17 in its multiplicative form:

$$Y_{uvt} = \exp\Bigg[\ln S_{vt} + \ln M_{ut} + \beta_1 \ln d_{uv} + \beta_2 \ln X_{uvt}\Bigg] + \epsilon_{uvt} \tag{1.23}$$

where Y_{uvt} is the bilateral dependent variable (credit flows, shares, number of clients served, etc.) aggregated at the urban-unit level. u, v and t are indices for origin (the urban unit from which the bank branches operate), destination (the urban unit in which the borrowing firms are located) and time. S_{vt} and M_{ut} are the origin urban unit $\times$ year and the destination urban unit $\times$ year fixed effects; fixed-effects ensure the theoretical restrictions implied by structural gravity are satisfied. X_{uvt} is a vector of time-varying pair characteristics (e.g. trade of goods between urban units u and v, dummy variable for belonging to the same department, region, etc.) that may affect firm-bank matching and financial decisions. Finally, I do not include a pair specific fixed effect as my goal is to identify the coefficient β_1 associated with physical distance d_{uv} between u and v. Here, the physical distance captures both the monitoring costs and the search frictions: according to my model and consistent to gravity estimation in international trade and finance, β_1 should be negative.

Impact of Broadband Internet Diffusion. I study the impact of a reduction of search frictions on aggregate credit flows using the staggered diffusion of Broadband Internet, i.e. prediction 1.18 of the model. I impose the following functional form for log $\kappa_{uvt} = \gamma\, \mathbb{C}_{uvt} + \varrho X_{uvt} + \varepsilon_{uvt}$. Inspired by the international trade literature that investigate the impact of time-varying trade policies as trade agreements, I consistently identify the effect of time-varying connectivity between urban units using a dynamic PPML estimator with fixed-effect in a difference-in-difference setting. Namely, I add to the baseline equation the time-varying variable of interest $\mathbb{I}_{uvt}$ that captures the pair specific variation in online connectivity (the reduction of search frictions):

$$Y_{uvt} = \exp\Bigg[\ln S_{ut} + \ln M_{vt} + \beta_1 \ln d_{uv} + \beta_2 \ln X_{uvt} + \beta_3\, \mathbb{C}_{uvt}\Bigg] + \epsilon_{uvt} \tag{1.24}$$

The goal is to consistently estimate the average effect of $\mathbb{C}_{uvt}$, a continuous variable indicating the degree of internet connectivity between u and v, using a structural gravity specification

the ML estimator derived from the Poisson distribution, however it does not require the dependant variable to be Poisson distributed (Fally, 2015)

derived from the model. The *origin* $\times$ *year* and *destination* $\times$ *year* fixed effects – S_{vt} and M_{ut} – are crucial as they absorb all the time-varying impacts of Broadband Internet which are not pair specific[18]. Another widely used specification includes dyadic fixed-effects, namely origin-destination-time FE, in order to absorb all time-invariant pair characteristics that may be correlated with the likelihood of being mutually connected. I propose an alternative version of specification 1.24 including pair-specific fixed-effect: a direct consequence is that I cannot estimate the coefficient β_1 relating to the physical distance d_{uv} in this specification.

PPML: Difference-in-Difference with Many Zeros? A key aspect of the empirical strategy is based on the performance of the PPML estimator with multiple high-dimensional fixed-effects. Santos Silva and Tenreyro (2011) shows that the PPML estimator is well behaved (and outperforms the OLS) when the dependent variable displays a large proportion of zeros, using a Monte-Carlo approach. In this paper, I extend and adapt their simulation exercise to the exact case of my empirical setting. Not only my estimating sample contains a vast majority of urban unit pairs that do not exchange credit over the entire period 1998-2005 (more than 95% of the credit shares equal zero) but also I have to deal with panel data and a difference-in-difference approach.

I present simulation evidence on the performance of the PPML estimator when the panel data is generated by a constant elasticity model, with (i) a large proportion of zeros, (ii) a time-varying shock and (iii) when all units are not simultaneously treated. In these simulations, the non-negative dependent variable Y_{uvt} is generated so that $\mathbb{P}(Y_{uvt} = 0)$ is substantial and $\mathbb{E}(Y_{uvt}|X_{uvt})$=exp($\beta\, X^{'}_{uvt}$). A the best of my knowledge, this is the first simulation evidence of the performance of the PPML estimator in this particular setting, that echoes a wide range of papers in international trade that investigate the impact of trade policies within a gravity framework.

Following Santos Silva and Tenreyro (2011), the dependent variable Y_{uvt} – which is the total credit granted by branches in v to firms in u – is generated as a finite mixture model of the form $Y_{uvt} = \sum_{j=1}^{m_{uvt}} z_{juvt}$, where m_{uvt} is the number of components of the mixture, and z_{juvt} a continuous random variable with support in $\mathbb{R}+$, distributed independently of m_{uvt}. This data generation process has a direct economic interpretation in my framework. m_{uvt} is the number of bank branches located in v that serve firms in u, and z_{juvt} is the amount

[18] Broadband Internet not only affects search frictions, but might also impact urban unit sizes, exogenous production costs or branch and firms productivity – in this setting, that does not bias the β_3 point estimate because of the urban units-time fixed-effects.

of credit that each those banks lent their clients located in u. Because m_{uvt} and z_{juvt} are independant, $\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt}) \times \mathbb{E}(z_{juvt}|X_{uvt})$. As in Santos Silva and Tenreyro (2011), z_{juvt} is obtained from a gamma distribution with mean 1 and variance 2, which is equivalent to a $\chi^2_{(1)}$ random variable. This implies, that conditionally on m_{uvt}, Y_{uvt} follows a $\chi^2_{m_{uvt}}$ and then $\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt})$. m_{uvt} will be generated as a negative-binomial random variable with conditional mean $\exp(\beta\, X'_{uvt})$ and a variance equal to $a\mathbb{E}(m_{uvt}|X_{uvt})$ + $b\mathbb{E}(m_{uvt}|X_{uvt})^2$. I propose the following functional form for $\mathbb{E}(Y_{uvt}|X_{uvt})$:

$$\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt}) = exp(\beta_0 + \beta_1\, x_{1uv} + \beta_2\, x_{2uvt} + \beta_3\, x_{3uvt})$$

where x_{1uv} is the product of the sum of two time invariant variables drawn from a standard normal x_{1u} and x_{1v}. x_{2uvt} is a time-varying variable equal to 1 with probability close to 0.4. Formally, I impose a dynamic structure for x_{2uvt} by introducing the underlying variable $w_{2uvt} = \gamma w_{2uvt-1} + \rho\epsilon_{uvt}$, with $\gamma = 1.05$, $\rho = 1.5$ and ϵ_{uvt} is drawn from a standard normal, such that $x_{2uvt} = \mathbb{I}(w_{2uvt} > 0.6)$. x_{3uvt} is a city pair specific treatment dummy variable that equals one after the time of treatment t and zero before. The city pair specific treatment time is drawn from a discrete uniform distribution over the support $[t_0, T]$. Formally, I generate city specific treatment date $x_{3u} \sim U[t_0, T]$ and define $x_{3uvt} = \mathbb{I}(t > \max(x_{3u}, x_{3v}))$. Finally, I impose $\beta_0 = 0$, $\beta_1 =$ -1, $\beta_2 = 1$ and $\beta_3 = 0.1$. This functional form has again a direct economic interpretation. The time-invariant variable x_{1uv} (which is symmetrical by definition, i.e., $x_{1uv} = x_{1vu}$) represents the distance between u and v, x_{2uvt} models the time-varying determinants of credit flows between cities and x_{3uvt} is analogous to my broadband internet interconnection shock. To complete my simulation setting, I need to define the conditional variance of m_{uvt} and Y_{uvt}. I follow Santos Silva and Tenreyro (2011) considering the quadratic specification $\mathbb{V}ar(m_{uvt}) = a\mathbb{E}(m_{uvt}|X_{uvt}) + b\mathbb{E}(m_{uvt}|X_{uvt})^2$ so that:

$$\mathbb{V}ar(Y_{uvt}|X_{uvt}) = (1 + 2a)\mathbb{E}(m_{uvt}|X_{uvt}) + 2b\mathbb{E}(m_{uvt}|X_{uvt})^2$$

Picking the value of a and b allows to generate a high probability of zeros and different heteroskedasticity patterns. Table 1.8 presents the results obtained with 1,000 replicas of the simulation procedure described here, in which the number of cities N is set to 200 and T equals 10. The estimation sample is therefore composed of 200 $\times$ 200 $\times$ 10 = 400,000 observations. In the top panel, a = 50 and b = 0 while a is set to 1 and b to 5 in the bottom panel. The table displays the point estimate and the standard errors obtained with the different estimators, namely PPML and GPML[19] with and without origin $\times$ year and

[19] GPML stands for Gamma Pseudo-Maximum Likelihood is a PPML-like estimation procedure in which

destination × year fixed-effects. These results confirm and extend the findings of Santos Silva and Tenreyro (2006, 2011), showing that both the PPML and the GPML estimators are well behaved in the two cases considered. In particular, the coefficient of interest β_3 is consistently estimated with two-way unit × fixed effects.These findings are an additional reason that justify my empirical approach and the validity of my estimation procedure.

5.2. Firm-Branch Matching

I then test how Broadband Internet diffusion affects the likelihood of firm-branch matching and the number of remote firms from u financed by a bank branch located in v. For this purpose, I leverage bank branch-level data in order to verify *Prediction 4* in a dynamic event-study approach, similar to Magouyres et al. (2019). I estimate a dynamic specification where I allow the effect on a branch b located in city u at year t, to vary with time-from-treatment.

The level of observation being a branch b located in city u, I'm able to discretize the treatment status by setting treatment status to 1 after the city experienced its highest increase in the predicted treatment variable $\hat{Z}_{it}^{\text{city-level}}$. Formally, I define the year of treatment as $t_{i0} = \text{argmax}_t\{\Delta\hat{Z}_{it}^{\text{city-level}}\}$ and discretized treatment status as $\mathbb{1}\{t \geq t_{i0}\}$. The year of treatment for each city is denoted t_{i0}. I index time-to-treatment with d (negative before treatment, positive after). The sample covers the years 1998 to 2005, and I restrict the set of observations where $d \in \{-6, -5, ..., +4, +5\}$. The main estimating equation is as follows:

$$Y_{b(u)t} = \sum_{\substack{d=-5 \\ d\neq -1}}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0u}\} + \mathbf{x}'_{ut}\delta + \alpha_{b(u)} + \psi_{r(u),t} + \varepsilon_{ut} \tag{1.25}$$

where $\alpha_{b(u)}$ and $\psi_{r(u),t}$ are fixed effects for the branch b located in city i and for the department (of the city)-year, and $\mathbf{x}'_{ut}$ is a vector of time-dependent city-level covariates. I drop two indicator variables for $d = -5$ and $d = -1$. That restriction is necessary to avoid multi-colinearity and to identify the fully-dynamic underlying data generating process in the staggered design (Borusyak and Jaravel, 2017; Gross et al., 2018). To ensure that this restriction is not influential in the results, I display results with alternative normalizations in the robustness section.

The specification presented in equation (1.25) includes leads and lags. The inclusion of leads

the dependant variable is a share instead of the value in level.

allows us to assess the presence of pre-trends. We also estimate a simpler "semi-dynamic" specification where only the lags of the treatment are included, as presented in equation (1.26):

$$Y_{b(u)t} = \sum_{d=0}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0u}\} + \mathbf{x}'_{ut}\delta + \alpha_v + \psi_{r(u),t} + \varepsilon_{ut} \tag{1.26}$$

The event-study coefficients $\widehat{\beta}_d$ estimated from equation (1.25) can be interpreted causally under the identifying assumption that, conditional on receiving broadband over the period considered and conditional on bank branch and city fixed-effects, the *timing* of broadband roll-out is unrelated to the outcome. The presence of systematic local factors that would drive both broadband and trade would be cause for concern. This potential issue is investigated by assessing the sensitivity of the coefficients to the inclusion of a large set of controls and fixed effects meant to account for city and well as local labor market shocks. Finally, the outcome variable $Y_{b(u)t}$ measures several aspects of the branch lending activity: i) the average distance to clients, ii) the share of remote clients (located outside the branch's urban unit), and iii) the share of credit granted remotely (i.e. to remote clients).

6. Results

This section presents the results. I first show that inter-regional credit flows follow a gravity equation by estimating equation (1.23), using a pseudo-Poisson maximum likelihood approach. Then, I estimate equation (1.24) and document an increase in inter-regional credit flows triggered by the staggered roll-out of Broadband Internet, associated with a reduction in search frictions. Finally, I estimate equation (1.25) to document that Broadband Internet diffusion allows banks to match with new firms located remotely. I show the results for different specifications and assess the robustness of the results.

Gravity equation for aggregate trade flows. Table 1.3 shows the results for the estimation of equation (1.23), using a pseudo-Poisson maximum likelihood approach described in section 5.1. The dependant variable is the credit shares Π_{uvt}, defined as the amount of credit granted by bank branches located in u to SMEs located in v, over the total credit stock of SMEs in v. Column (1) display the results with no control, origin urban unit × year and destination urban unit × year fixed-effects. In columns (2) to (5), I sequentially add pair-level controls: a dummy variable equal to 1 if both the firm and the bank are located in the same region (*région*), in the same county (*départment*) and the lagged log of trade

flows between counties. The coefficient of interest β_1, associated to the log of the distance between urban unit u and v, is negative and close to -2. This magnitude implies that firms borrow credit from banks located in a close-by urban unit nearly four times more than from similar banks located at twice the distance. The distance coefficient decreases but remains close to -2 when I consider banks and firms located within the same region or county, or when I control by lagged trade flows in column (4) which is the baseline specification.

Table 1.3: Gravity Equation for Inter-regional Credit Flows

	Share of credit in v borrowed from u: Π_{uvt}				
	(1)	(2)	(3)	(4)	(5)
Log distance d_{uv}	-2.197***	-1.966***	-1.829***	-1.765***	-1.741***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Same region		1.482***			
		(0.018)			
Same county			1.722***		0.803***
			(0.016)		(0.023)
log (Trade Flows)				0.473***	0.304***
				(0.004)	(0.006)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
Observations	24,545,811	24,545,811	24,545,811	24,545,640	24,545,640

Notes: PPML estimation of equation (1.23). d_{uv} = bilateral distance. Same region is a dummy variable equal to 1 if both the firm and the bank are located in the same region. Same county is a dummy variable equal to 1 if both the firm and the bank are located in the same county. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit. Robust standard errors in in parentheses.

It is not evident to compare this magnitude with the literature: as far as I know, this paper is the first to estimate gravity equations for inter-regional credit flows. Portes and Rey (2005) document the negative impact of distance for international cross-border equity flows, and find coefficients twice as low, between -0.529 and -0.881. Recently, Brei and von Peter (2018) run a similar estimation on international banking flows and adopt a PPML approach. They find an estimate close to but smaller than one. A coefficient of magnitude -2 is also larger than what is documented by the vast trade literature, comprising more than 2,500 estimates of the distance effect (Head and Mayer, 2014). Therefore, my results indicates that

distance-related frictions are likely to be more important at a very local level than for international credit flows. This confirm the local nature of credit markets for SMEs and echoes the first-order importance of distance as a determinant of access to credit, well documented in the banking literature (e.g., Petersen and Rajan, 1995; Agarwal and Hauswald, 2010)

For all $u - v$ pairs with no bilateral credit flow, the share of credit is set to zero, which represents the vast majority of the observations. For robustness, I run a similar estimation but keeping only pairs of urban units with positive credit flows. Table 1.5 shows the results. The coefficient of interest is still negative and significant, but of a smaller magnitude: β_1 is now close to -1 and remains stable to the addition of control variables.

Impact of a reduction in search frictions. I now test how the gravity equation for inter-regional credit flows is distorted by a technology-induced reduction in search frictions. I formally test the prediction (1.18) of my model by running a PPML estimation of the augmented gravity equation (1.24). The main results of my paper are presented in Table 1.4, where the first column reproduces column (4) of Table 1.3 for comparison purposes. In column (2), I include the continuous variable $\mathbb{C}_{uvt}$ for observed internet interconnection between two urban units. Similar to Table 1.3, in order to estimate an effect on distance, origin and destination $\times$ year fixed effects are included but not pair specific fixed effects. The estimate is positive and significant, in line with the model prediction. It suggests that a technology-induced reduction in search frictions distort the gravity equation for credit flows. The amount of credit exchanged between firms and banks located in connected urban units increases relative to other not-connected banking partners. The magnitude of the estimate implies that the share of credit granted to firms located in v from banks located in a remote urban unit u increases by 44% on average when u and v are connected. The estimate for distance is not affected by the inclusion of the ADSL variable, nor the control for bilateral trade flows.

My model predicts an heterogeneous effect of brandband internet with respect to distance which is the result of two opposing forces: a connectivity effect (positive impact) and a competition effect (negative impact). While the former dominates overall, the competition effect could prevail for markets that are geographically close, as banks located in these close markets already benefit from a visibility advantage. In contrast, remote submarkets benefit more for a reduction of search frictions. I formally test this prediction by including the interaction variable for the bilateral distance (expressed as deviation from the sample average)

and the treatment variable in column (3). I find a positive and statistically significant effect of the interaction variable with distance, which means that the elasticity of credit flows with respect to distance decreases in magnitude with broadband internet. In other words, the positive impact of a reduction in search frictions on credit flows is higher when two very distant cities are connected, which verifies the intuition of the model. On the contrary, the effect is almost null or even negative when two neighbouring cities, already economically very closely tied, are interconnected by internet.

Table 1.4: Technology-Induced reduction in search frictions

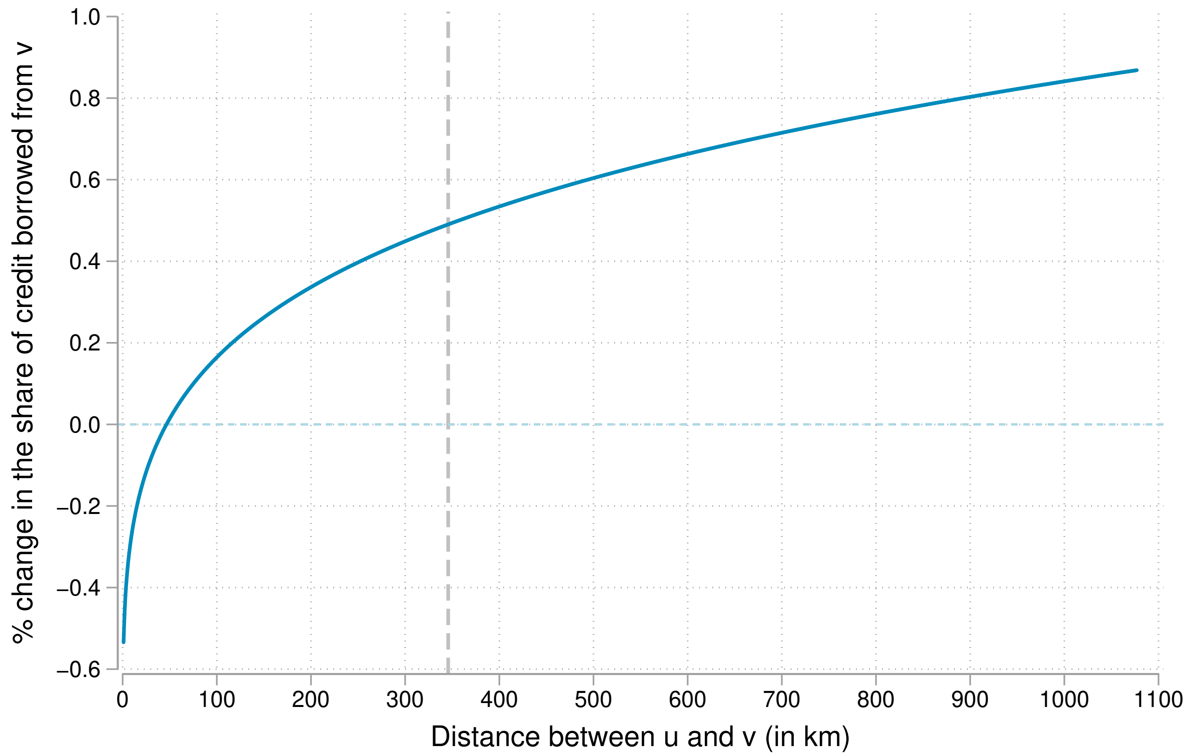
	Share of credit in v borrowed from u: Π_{uvt}				
	(1)	(2)	(3)	(4)	(5)
Log distance d_{uv}	-1.765***	-1.760***	-1.844***	-1.764***	-1.865***
	(0.004)	(0.004)	(0.006)	(0.004)	(0.006)
$\mathbb{C}_{uvt}$		0.370***	0.800***		
		(0.058)	(0.059)		
Log distance d_{uv} x $\mathbb{C}_{uvt}$			0.230***		
			(0.008)		
$\hat{\mathbb{C}}_{uvt}$				0.058	0.399***
				(0.059)	(0.062)
Log distance d_{uv} x $\hat{\mathbb{C}}_{uvt}$					0.199***
					(0.008)
log (Trade Flows)	0.473***	0.475***	0.470***	0.474***	0.468***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
Observations	24,545,640	24,545,640	24,545,640	24,545,640	24,545,640

Notes: PPML estimation of equation (1.23). d_{uv} = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$, is a continuous variable that indicates the degree of broadband internet inter-connectivity between two urban units. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My instrument variable strategy delivers a similar measure of between urban unit connectivity named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit. Robust standard errors in in parentheses.

An important concern with those results is that potential endogeneity of internet take-up biases these estimates. In columns (4) and (5), I estimate equation (1.24) using the instrument variable for broadband internet interconnection, namely $\hat{\mathbb{C}}_{uvt}$, instead of the observed connection variable used in columns (2) and (3). The sign of the coefficient of interest remains

unchanged, although the magnitude of the effect declines. The overall effect in column (4) is imprecisely estimated and implies that the share of credit granted to firms located in v from banks located in a remote urban unit u now increases by only 6% when u and v are connected.

Fig. 1.1. Heterogeneity of the effect with respect to distance



Notes: This figure plots the marginal effect of broadband internet with respect to distance between to interconnected cities u and v, as estimated in equation (1.24). The x-axis represents the distance in kilometers and the y-axis shows the total effect of broadband internet connection on the share of credit borrowed from v by firms located in u, in %. The vertical grey dashed line represent the average distance between two cities in my sample, and correspond to the overall effect estimated in column (4) of Table 1.4.

The results in column (5) documenting the heterogeneity of the effect with respect to distance are very comparable to those in column (3) that do not use the instrumental variable. In particular, the interaction term is positive and significant. In economic terms, theses results mean that an increase in internet availability of 10 percentage points increases credit flows for an urban unit at the 25th distance percentile by 24% less than for an urban unit at the 75th distance percentile. Figure 1.1 illustrates this heterogeneity. While the effect of being interconnected is negative when cities are geographically nearby (competition effect dominates), it increases sharply and because positive after the 50th kilometer. After the 100th kilometer the slope of the curve is then much flatter. Figure 1.15 in Appendix maps

the heterogeneous effect of being connected to Paris, showing that Marseille (second biggest French city) benefit more than Lyon (third biggest city).

I then test for robustness along different dimensions. I first estimate the augmented gravity equation (1.18) by adding bilateral pair fixed-effects that control for any unobserved characteristics of the urban unit pair that are constant over time (see Head and Mayer, 2014). Table 1.6 shows the results in columns (2) and (4). Although less significant and of lower magnitude, they are virtually unchanged and confirm the positive but heterogeneous effect of the technology-induced reduction in search frictions. Second, in order to take into account the dynamic nature of credit flows, I replicate the baseline results by adding the lag dependant variable to the regressors in Table 1.7. This implies that credit relationships existing in the previous year provides a basis for the credit flows observed in the current year.[20] The main findings are confirmed by this alternative specification.

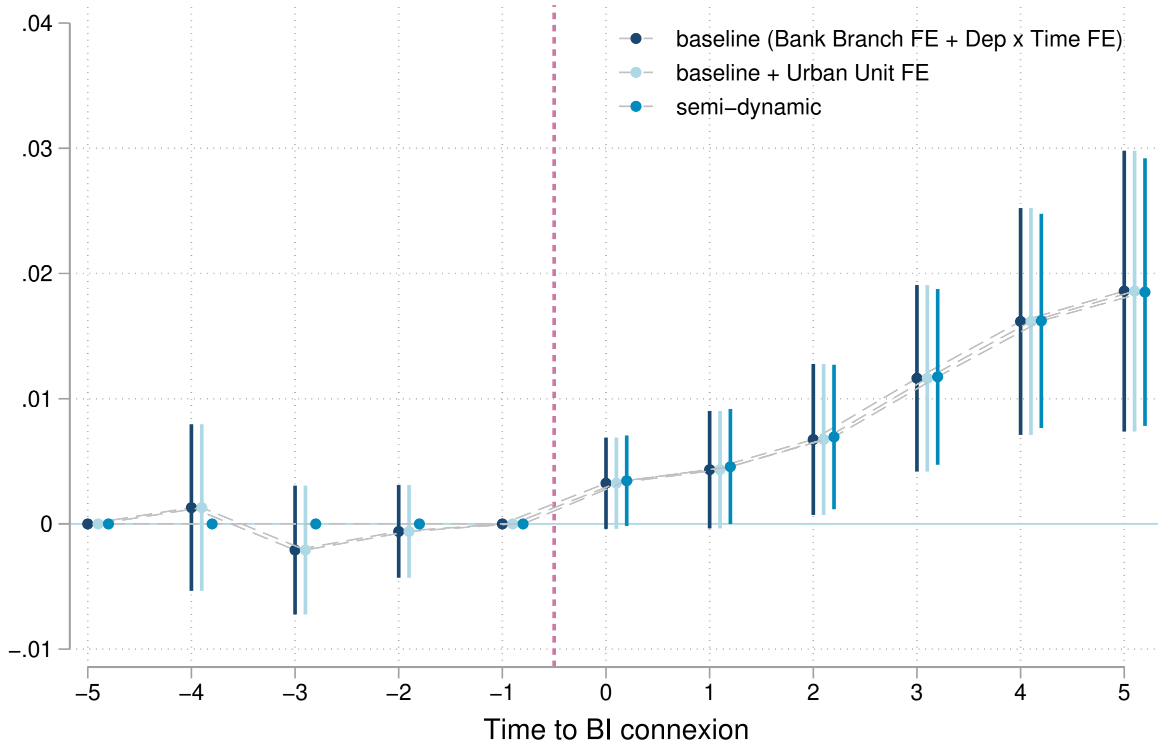
Firm-bank matching. The model predicts that broadband internet diffusion affects the firm-branch matching process: the reduction in search frictions allows very distant firms and bank to match and, in turn, the share of remote firms (located outside u) financed by a bank located in u increases. I formally test prediction (1.22) leveraging micro bank branch-level data in the dynamic event-study setting described above in section 5.2.

The main variable I consider is the share of credit lent by bank branches located in a given city to firms located outside that city. With a high level of search frictions, matches between firms and banks occur only locally, within cities, and the share of remote credit is close to zero. On the contrary, if search is free, firms can meet with and borrow from banks located far away and the share of credit lent remotely is large. I show the results for different specifications and assess the robustness of the results. I then turn to the extensive margin: the share of firms located remotely financed by a bank in a given city.

Figure 1.2 displays the results, plotting estimated coefficients from equation (1.25). The dark blue dots report results from the baseline specification including the bank branch and county-year fixed-effects. Estimates exhibit a flat trend before the event (i.e. the normalizing measure of time since access d = -1) and a break in the trend after that. The coefficient for d = 5 in that specification is 0.018 suggesting that the expansion of access to broadband

[20]However, given that the time dimension is much lower than N, estimates are likely to suffer from the Nickell bias in the dynamic model.

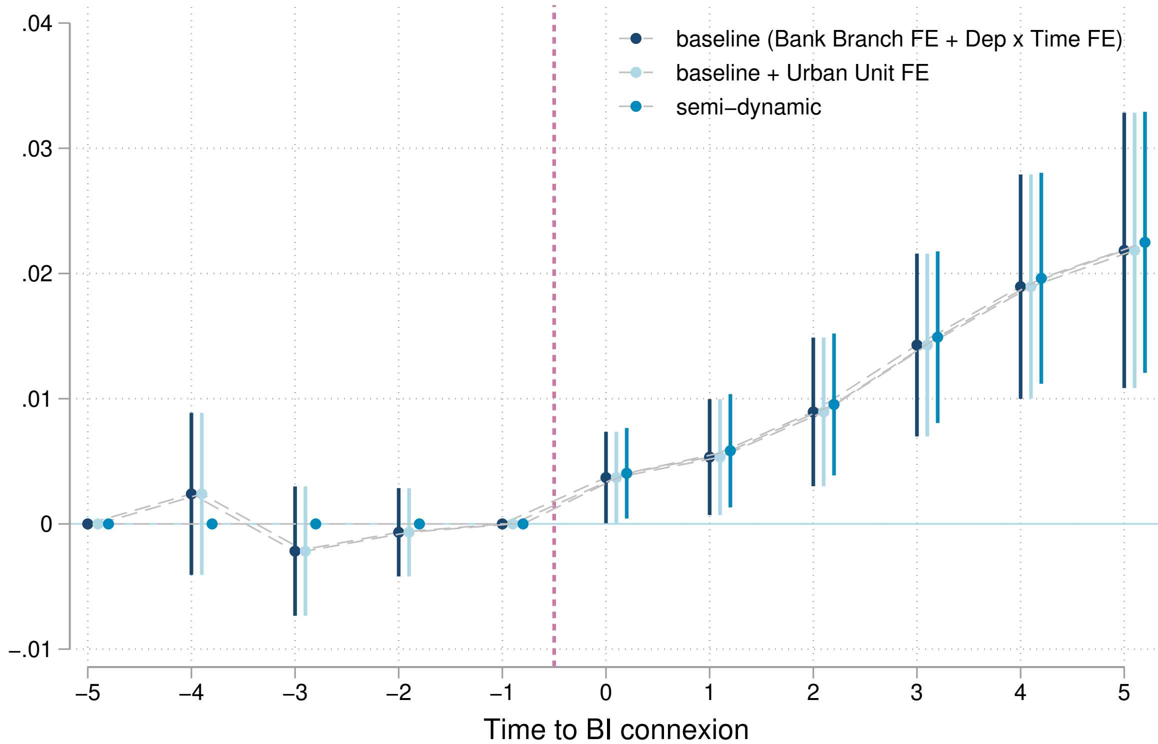
Fig. 1.2. Share of credit to remote firms



Notes: This figure plots estimates for specification in equation (1.25) – fully dynamic – and equation (1.26) – semi-dynamic. The dependant variable is the share of credit granted to firms located outside the bank's urban unit. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

internet increased the bank branch-level share of credit lent to distant firms by about 10%, 5 years after the period of largest expansion. Not only this is in line with the model predictions but also it is consistent with the aggregate impact on credit flows documented above. Our second specification adds urban unit fixed-effects to the regression that aim to control for time invariant city characteristics. The light blue dot show the estimates. Here again, I find no sign of a pre-trend prior to broadband expansion contrasting with a steady growth afterwards. The estimated effect after five years is virtually the same than in the baseline case. Finally, the last set of coefficients plotted in royal blue represent a semi-dynamic version of the baseline specification (see Equation 1.26). The regression should in theory more efficiently estimated –as the number of parameters to be estimated is lower–, however both the estimates and the standard errors stay very stable and close to the fully dynamic specification in practice.

Fig. 1.3. Share of remote clients



Notes: This figure plots estimates for specification in equation (1.25) – fully dynamic – and equation (1.26) – semi-dynamic. The dependant variable is the share of clients financed outside the bank's urban unit. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

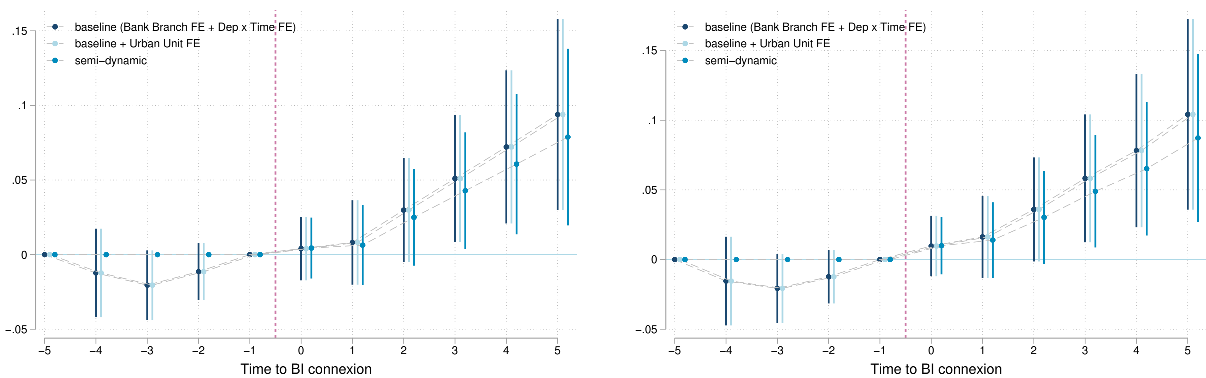
I then explore the possible mechanisms behind the bank-level increase in share of credit lent to remote clients documented above. I study how the extensive margin of credit, i.e, the number of banking relationships, is affected by the technology-induced reduction in search frictions (as opposed to the average credit, the intensive margin). For this purpose, the dependant variable I consider is the share of remote firms financed by bank branches located in a given city, which is defined as the number of clients located outside the bank's urban unit divided by the total number of clients. Figure 1.3 displays the results. Similar to the previous estimation, I find a flat pre-trend and a positive effect afterwards. The coefficient for d = 5 now equals 0.023, implying that the technology-induced reduction in search frictions increased the bank branch-level share of remote clients by about 12%, 5 years after the period of largest expansion. Interestingly, the effect on the extensive margin is comparable but slightly higher than the overall effect. This suggests that (i) the increase in between cities credit flows is mainly driven by the creation of new relationships (new matches be-

tween banks and firms located in different cities), and (ii) that those new credit relationships are in average smaller than the existing ones.

A direct consequence of these results is that firm-bank distance should increase following Broadband Internet expansion, because there are more matches between firms and banks located in different cities. I directly verify this hypothesis by running a similar dynamic event-study regression with the (weighted) firm-bank distance as dependant variable. Figure 1.4 displays the results.

Fig. 1.4. Firm-bank distance

(a) Unweighted (b) Weighted by credit exposure



Notes: These figures plot estimates for specification in equation (1.25) – fully dynamic – and equation (1.26) – semi-dynamic. In the left panel, the dependant variable is the (log) average firm-bank distance measured at the branch level. In the right panel, the (log) distance is weighted by credit exposure. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

Figure 1.4(a) documents a positive and significant effect of Broadband Internet on firm-bank distance. The magnitude implies that connected banks match with firms that are located in average 10% further, 5 years after the shock. Figure 1.4(a) shows that the result is not significantly affected if one considers the average distance weighted by credit exposure. These results are in line with Kroszner and Strahan (1999) point out that innovation in information technology reduced the dependence on geographical proximity between customers and banks in the US, starting in the 70s. This is also consistent to Petersen and Rajan (2002) that documents the erosion of the local nature of small business lending, with increasing distance between small firms and their lenders in the United States but also new communication habits. Similar trends are observed in France: inter-regional credit flows have grown by 15%

and the average firm-bank distance has increased by 10% between 1998 and 2005. As far as I know, this paper is the first to provide a causal interpretation for those facts, suggesting that innovations in information technology – namely, Broadband Internet diffusion – reduced the role of both transaction and search costs in shaping credit outcomes, allowing firms to search for credit further and leading to structural changes in local credit markets. This suggests that innovations in information technology reduced the role of both transaction and search costs in shaping credit outcomes, allowing firms to search for credit further and leading to structural changes in local credit markets.

7. Implications for the cost of debt

In this last section, I use the empirical results from Section 6 to quantify the impact of the technology-induced reduction in search frictions on loan prices through the lens of my model. Mapping the model prediction (1.17) into the gravity equation I estimate gives the following equivalence between the model parameters and the empirical estimates $\beta_1 = \nu\,\theta$, $\beta_2 = \varrho$ and $\beta_3 = \gamma$. Thus, the technology-induced reduction in search frictions κ_{uv} for a pair of connected cities formally writes:

$$\Delta \ln\kappa_{uv} = \hat{\gamma}\,\Delta\,\mathbb{C}_{uv} = -0.058 \tag{1.27}$$

The distribution of the minimum loan price (Equation 1.9) obtained by an entrepreneur located in v rewrites as follow:

$$\begin{aligned}\mathbb{P}\big(r_{e_v} \leq r\big) = W_{e_v}(r) &= 1 - \exp\Bigg(-r^{\theta} z_{e_v}^{\theta+1} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv}\Bigg) \\ &= 1 - \exp - \Bigg(\frac{r}{z_{e_v}^{-\frac{\theta+1}{\theta}} (\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv})^{-\frac{1}{\theta}}}\Bigg)^{\theta}\end{aligned}$$

Which delivers the following expected lowest rate P_{e_v} for the entrepreneur e_v:

$$P_{e_v} = \mathbb{E}[r|e_v] = z_{e_v}^{-\frac{\theta+1}{\theta}} (\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv})^{-\frac{1}{\theta}} \times \Gamma\big[1 + \frac{1}{\theta}\big] \tag{1.28}$$

where Γ stands for the Gamma function. By taking the log of Equation 1.28, I can isolate the effect on P_{e_v} of the technology-induced reduction in search frictions caused by the diffusion

of Broadband Internet:

$$\Delta ln P_{e_v} = ln\, \frac{P_{e_v}(1)}{P_{e_v}(0)} = -\frac{1}{\theta}\, ln \left(\frac{\sum_{u=1}^{N} \hat{\kappa}_{uv}(1) \cdot S_u(0) \cdot (c_u d_{uv})^{-\theta}(0)}{\sum_{u=1}^{N} \hat{\kappa}_{uv}(0) \cdot S_u(0) \cdot (c_u d_{uv})^{-\theta}(0)} \right) \tag{1.29}$$

where (1) indicates the situation after Broadband Internet access and (0) before. As the origin fixed-effect FE_u equals $S_u \cdot c_u^{-\theta}$ and $d_{uv}^{-\theta}$ equals $dist_{uv}^{-\nu\theta}$, equation (1.29) simplifies as follow:

$$\Delta ln \hat{P_{e_v}} = -\frac{1}{\theta}\, ln \left(\frac{\sum_{u=1}^{N} \hat{\kappa}_{uv}(1) \cdot F\hat{E}_u(0) \cdot dist_{uv}^{-\hat{\nu}\hat{\theta}}}{\sum_{u=1}^{N} \hat{\kappa}_{uv}(0) \cdot F\hat{E}_u(0) \cdot dist_{uv}^{-\hat{\nu}\hat{\theta}}} \right) \tag{1.30}$$

While Broadband Internet may affect all variables in this equation as the firm productivity z_{e_v}, the cost c_u and the θ paramater of the bank branches' size distribution, I keep them constant to conduct counterfactual exercises. I finally plug into this equation my empirical estimates for $\beta_1 = -\nu\theta$, FE_u and $\kappa_{uv}(1) = \gamma$ as well as parameters calibrated from the data (θ, $dist_{uv}$) in order to compute $\Delta ln \hat{P_{e_v}}$ the change in the lower expected cost of debt triggered by the BI-induced reduction in search frictions. My model predicts an average decline of -4.9% in 2005, compared to what it would have been without any lowering of search and contracting costs. This decline in the cost of debt would have been higher if all french cities were connected at the end of 2005, with an average value of -5.8%. This results echoes the conclusion of Hauswald and Marquez (2003) that shows how improved access to information makes markets more competitive so that customers benefit from technological progress. The mechanism herein departs from theirs as I focus on firm search and I do not model the process by which banks search for customers (and information about those customers), but the intuition of an easier *dissemination of information* is similar.

Figure 1.13 illustrates the spatial dimension of the decline in firms' cost of debt. Dark blue cities are the one in which this decline is the strongest (larger than the 75^{th} percentile) while the light grey ones indicates a reduction lower than the median. Figure 1.14 gives a similar point of view with a focus on the Paris region. It is noteworthy that all the largest french cities (Paris, Marseille, Bordeaux or Lyon) benefit less from this decline than suburban or rural areas. This results is in line with the model intuition that a reduction in search frictions precipitate an increase in competition due to the ability for firm to search further and meet with more bank branches. In already crowded markets with a lot of active banks and a high level of competition – typically in large city centers – firms were not highly constrained by the search frictions and the decline is low. On the contrary, in isolated and

rural submarkets where firms had to make costly efforts to multiply meetings with different bankers and eventually match with the right one, the reduction in search frictions triggers a substantial decline in loan prices.

8. Conclusion

I develop a new theory of firm-bank matching subject to search frictions. I provide a causal evidence of such frictions affect firm-bank matching and the allocation of bank credit, using the staggered roll-out of broadband internet in France as a shock on transaction and search costs. I show that this technology-induced reduction in search frictions triggers an increase by 6% of the share of credit exchanged between interconnected cities. This positive effect varies with the initial level of search frictions: it is higher when two very distant cities are connected. On the contrary, the effect the effect is almost null when two neighboring cities, already economically very closely tied, are interconnected by internet. Leveraging bank branch-level data, I document that Broadband Internet diffusion allows banks to match with new firms located remotely. Connected banks increase their share of credit lent to firms located outside their city by 10%, and their share of remote clients by almost 12%. As a result, the average distance between a bank and its customers increases by 10% in the medium run after broadband internet access. Finally, I plug these estimates into the equation linking search frictions to loan prices. Interpreted within my model, the reduced-form estimates imply that the reduction in search frictions due to the large diffusion of Broadband Internet lowered the cost of debt for small businesses by 4.9% on average. Overall, this paper highlight the role of transactions and search cost in shaping firm's access to credit. Credit markets with high search frictions make financing by bank credit both difficult, time-consuming and onerous, especially for small businesses. This conclusion calls for a variety of economic policies aiming at to make the process of searching and applying for credit more fluid, efficient and less burdensome, in particular in a period of pandemic marked by the disappearance of face-to-face interactions and the consequent surge of digitalization

References

Agarwal, S. and Hauswald, R. (2010). Distance and Private Information in Lending. *The Review of Financial Studies*, 23(7):2757–2788.

Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Akerman, A., Leuven, E., and Mogstad, M. (2018). Information frictions, internet and the relationship between distance and trade. *mimeo*, page 48.

Allen, J., Clark, R., and Houde, J.-F. (2019). Search frictions and market power in negotiated-price markets. *Journal of Political Economy*, 127(4):1550–1598.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Arcep (2002). L'accés haut débit via l'adsl : historique des décisions de l'art et de l'arcep. Technical report.

Argyle, B., Nadauld, T., and Palmer, C. (2019). Real effects of search frictions in consumer credit markets. *MIT Sloan School Working Paper 5242-17*.

Berger, A., Leusner, J. H., and Mingo, J. J. (1997). The efficiency of bank branches. *Journal of Monetary Economics*, 40(1):141–162.

Berger, A., Miller, N. H., Petersen, M., Rajan, R., and Stein, J. (2005). Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial Economics*, 76(2):237–269.

Berger, A. and Udell, G. (1995). Relationship lending and lines of credit in small firm finance. *The Journal of Business*, 68(3):351–81.

Berger, A. and Udell, G. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *Economic Journal*, 112(477):F32–F53.

Bhuller, M., Kostøl, A., and Vigtel, T. C. (2019). How Broadband Internet Affects Labor Market Matching. Memorandum 10/2019, Oslo University, Department of Economics.

Boot, A. W. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, 9(1):7–25.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs with an application to the estimation of the marginal propensity to consume. *mimeo*, page 33.

Boualam, Y. M. (2018). Credit markets and relationship capital. *Working Paper at University of Pennsylvania,*.

Brei, M. and von Peter, G. (2018). The distance effect in banking and trade. *Journal of International Money and Finance*, 81:116–137.

Cerqueiro, G., Degryse, H., and Ongena, S. (2011). Rules versus discretion in loan rate setting. *Journal of Financial Intermediation*, 20(4):503 – 529.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634.

Coeurdacier, N. and Martin, P. (2009). The geography of asset trade and the euro: Insiders and outsiders. *Journal of the Japanese and International Economies*, 23(2):90–113.

Cole, R., Goldberg, L. G., and White, L. (2004). Cookie cutter vs. character: The micro structure of small business lending by large and small banks. *Journal of Financial and Quantitative Analysis*, 39(02):227–251.

Correia, S., Guimaraes, P., and Zylkin, T. (2019). ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects. *arXiv e-prints*, page arXiv:1903.01690.

Degryse, H., Kim, M., and Ongena, S. (2009). *Microeconometrics of Banking Methods, Applications, and Results*. Oxford University Press.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *The Journal of Finance*, 60(1):231–266.

den Haan, W. J., Ramey, G., and Watson, J. (2003). Liquidity flows and fragility of business enterprises. *Journal of Monetary Economics*, 50(6):1215–1241.

Drexler, A. and Schoar, A. (2014). Do relationships matter? evidence from loan officer turnover. *Management Science*, 60(11):2722–2736.

Duquerroy, A., Mazet-Sonilhac, C., Mésonnier, J.-S., and Paravisini, D. (2019). Branch closures and specialization. *Mimeo*.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Eaton, J., Kortum, S., and Kramarz, F. (2018). Firm-to-Firm Trade: Imports, exports, and the labor market. Discussion papers 16048, Research Institute of Economy, Trade and Industry (RIETI).

Ernst and Young (2018). The future of sme banking. *EY Financial Services Report*.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, 97(1):76–85.

FED (2014). Small business credit survey.

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1):185–206.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica*, 52(3):681–700.

Gross, T., Notowidigdo, M. J., and Wang, J. (2018). The marginal propensity to consume over the business cycle.

Hauswald, R. and Marquez, R. (2003). Information Technology and Financial Services Competition. *The Review of Financial Studies*, 16(3):921–948.

Hauswald, R. and Marquez, R. (2006). Competition and Strategic Information Acquisition in Credit Markets. *The Review of Financial Studies*, 19(3):967–1000.

Head, K. and Mayer, T. (2014). Chapter 3 - gravity equations: Workhorse,toolkit, and cookbook. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, volume 4 of *Handbook of International Economics*, pages 131 – 195. Elsevier.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153):41–57.

Hubbard, R., Kuttner, K., and Palia, D. N. (2002). Are there bank effects in borrowers' costs of funds? evidence from a matched sample of borrowers and banks. *The Journal of Business*, 75(4):559–81.

Infosys (2018). Banks! it's time to change your game in smes lending.

Ioannidou, V. and Ongena, S. (2010). "time for a change": Loan conditions and bank behavior when firms switch banks. *Journal of Finance*, 65(5):1847–1877.

Jing, H. (2014). Financial strength and monitoring capital: Evidence from firm-bank matching. *Working paper, Xiamen University.*

Kroszner, R. S. and Strahan, P. E. (1999). What drives deregulation? economics and politics of the relaxation of bank branching restrictions. *The Quarterly Journal of Economics*, 114(4):1437–1467.

Lendle, A., Olarreaga, M., Schropp, S., and Vézina, P.-L. (2016). There goes gravity: ebay and the death of distance. *Economic Journal*, 126(591):406–441.

Lenoir, C., Mejean, I., and Martin, J. (2018). Search Frictions in International Good Markets. Technical report.

Magouyres, C., Mayer, T., and Mazet-Sonilhac, C. (2019). Technology-Induced Trade Shocks? Evidence from Broadband Expansion in France. (2019-10).

Malgouyres, C. (2017). The impact of chinese import competition on the local structure of employment and wages: Evidence from france. *Journal of Regional Science*, 57(3):411–441.

McAllister, P. H. and McManus, D. (1993). Resolving the scale efficiency puzzle in banking. *Journal of Banking & Finance*, 17(2-3):389–405.

McKinsey (2018). The lending revolution: How digital credit is changing banks from the inside. *McKinsey Report.*

Nguyen, H.-L. Q. (2019). Are Credit Markets Still Local? Evidence from Bank Branch Closings. *American Economic Journal: Applied Economics*, 11(1):1–32.

OECD (2018). Enhancing sme access to diversified financing instruments.

Okawa, Y. and van Wincoop, E. (2012). Gravity in international finance. *Journal of International Economics*, 87(2):205 – 215.

Paravisini, D., Rappoport, V., and Schnabl, P. (2015). Specialization in bank lending: Evidence from exporting firms. Working Paper 21800, National Bureau of Economic Research.

Petersen, M. and Rajan, R. (2002). Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570.

Petersen, M. A. and Rajan, R. (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443.

Portes, R. and Rey, H. (2005). The determinants of cross-border equity flows. *Journal of International Economics*, 65(2):269 – 296.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of finance*, 47(4):1367–1400.

Rauch, J. E. (2001). Business and social networks in international trade. *Journal of Economic Literature*, 39(4):1177–1203.

Rogerson, R., Shimer, R., and Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988.

Salop, S. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10(1):141–156.

Santos Silva, J. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Santos Silva, J. and Tenreyro, S. (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112(2):220–222.

Schwert, M. (2018). Bank capital and lending relationships. *Journal of Finance*, 73(2):787–830.

Sénat (2002). Le bilan de la loi n 96-659 de réglementation des télécommunications. *Rapport d'information du Sénat*.

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The Journal of Finance*, 45(4):1069–1087.

Stein, J. C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *The Journal of Finance*, 57(5):1891–1921.

Steinwender, C. (2018). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–96.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410.

Telecom, F. (2003). Conférence de presse du 10 juin 2003 "internet haut débit pour tous : France télécom s'engage". Technical report.

Tregouet, R. (2001). Question écrite *n*. 30844 (sénat, 01/02/2001). *Senat*.

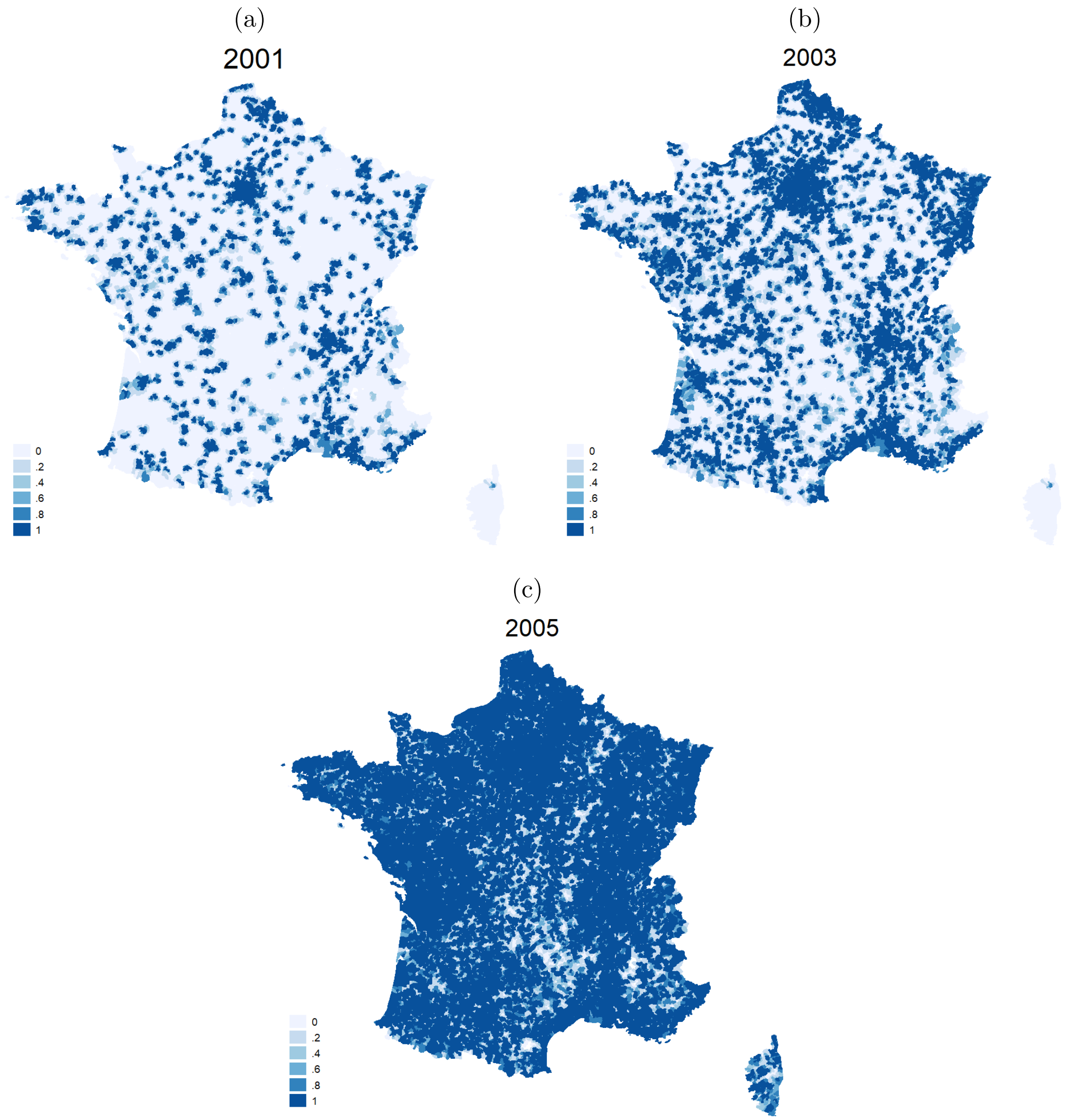
Udell, G. F. (2015). SME Access to Intermediated Credit: What Do We Know and What Don't We Know? In Moore, A. and Simon, J., editors, *Small Business Conditions and Finance*, RBA Annual Conference Volume (Discontinued). Reserve Bank of Australia.

Wasmer, E. and Weil, P. (2004). The macroeconomics of labor and credit market imperfections. *American Economic Review*, 94(4):944–963.

Wheelock, D. and Wilson, P. (2012). Do large banks have lower costs? new estimates of returns to scale for u.s. banks. *Journal of Money, Credit and Banking*, 44(1):171–199.

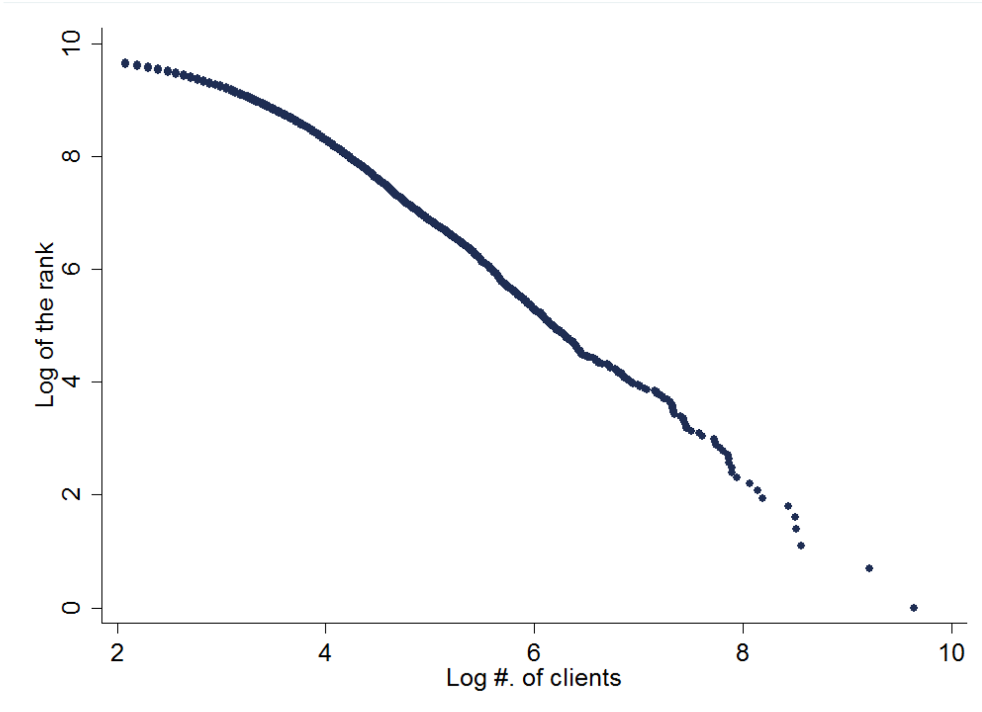
9. Figures

Fig. 1.5. Broadband internet roll-out in France



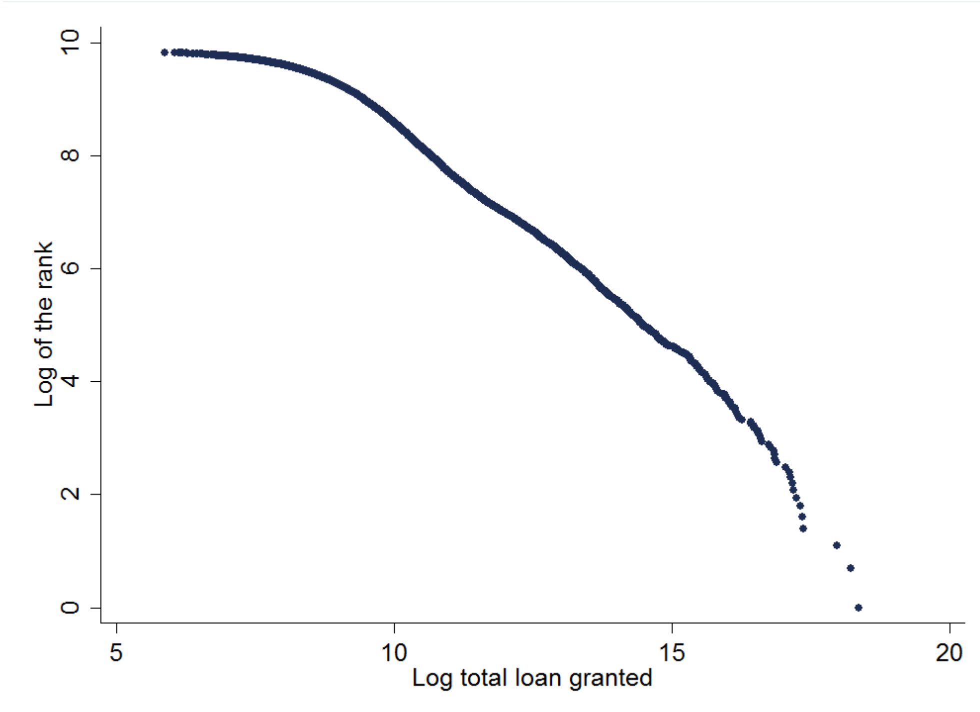
Notes: This figure shows the roll-out of Broadband Internet for all city in mainland France, for the years 2001, 2003 and 2003. The dark blue areas represent a large degree of coverage (Z_{ut} close to 1), while the light blue areas are city with no internet connection (Z_{ut} equal to 0). This unique data was collected by Magouyres et al. (2019) and contain the date of upgrade to ADSL for each Local Exchange (LE)'s in mainland France. The historical operator (France Télécom) had to make this data available to other operators as well as websites allowing consumers to gauge the quality of their line for regulatory reasons. Additionally, the authors gather data from the regulatory agency (ARCEP) regarding the geographical coverage of each LE. Combining both datasets,they construct a continuous measure Z_{ut} described in Section 1.

Fig. 1.6. Branch Rank versus Size (Total credit)



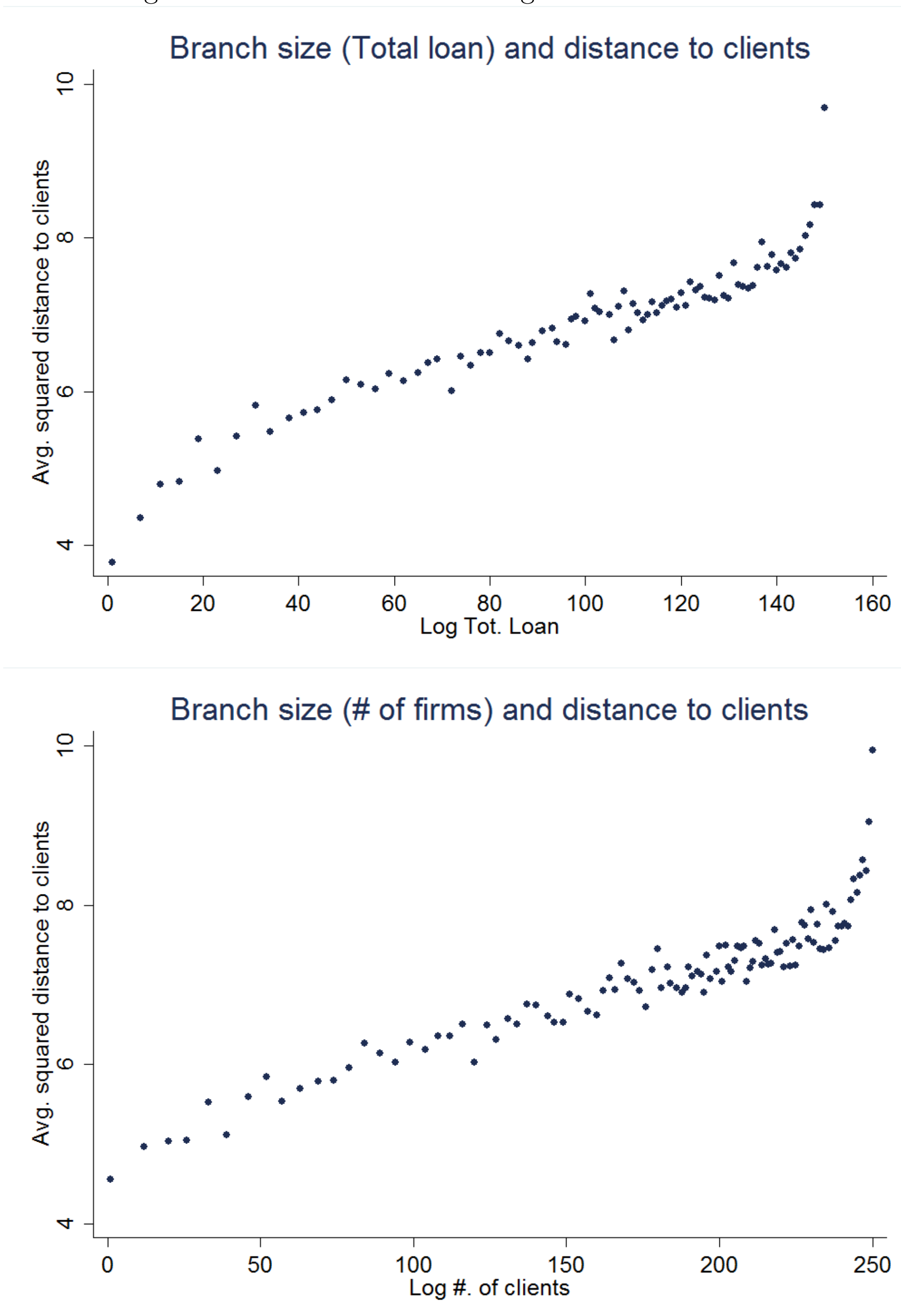
Notes: This figure displays the distribution of bank branch size for the last quarter of 2005, for all banks with at least 5 clients. Formally, it shows scatter plot of the log (Size) against the log (Rank). I compute the size of a branch as its total credit exposure. Bank branches are ranked by size: #1 being the largest branch, #2 the second largest, and so on.

Fig. 1.7. Branch Rank versus Size (#. clients)



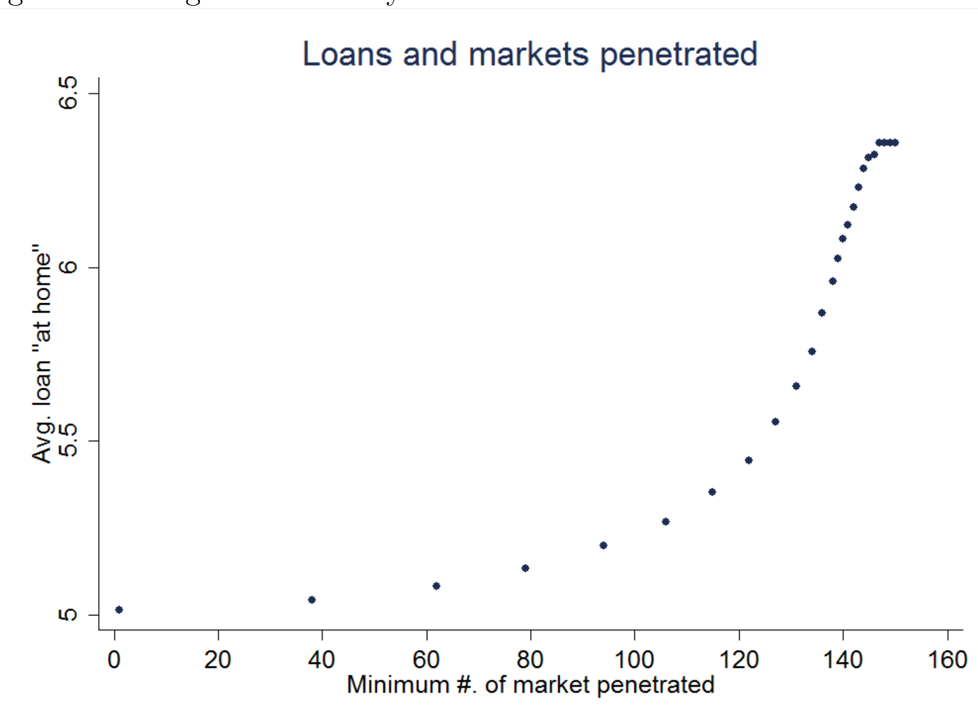
Notes: This figure displays the distribution of bank branch size for the last quarter of 2005, for all banks with at least 5 clients. Formally, it shows scatter plot of the log (Size) against the log (Rank). I compute the size of a branch as its number of clients. Bank branches are ranked by size: #1 being the largest branch, #2 the second largest, and so on.

Fig. 1.8. Branch Size and Average Distance to Clients



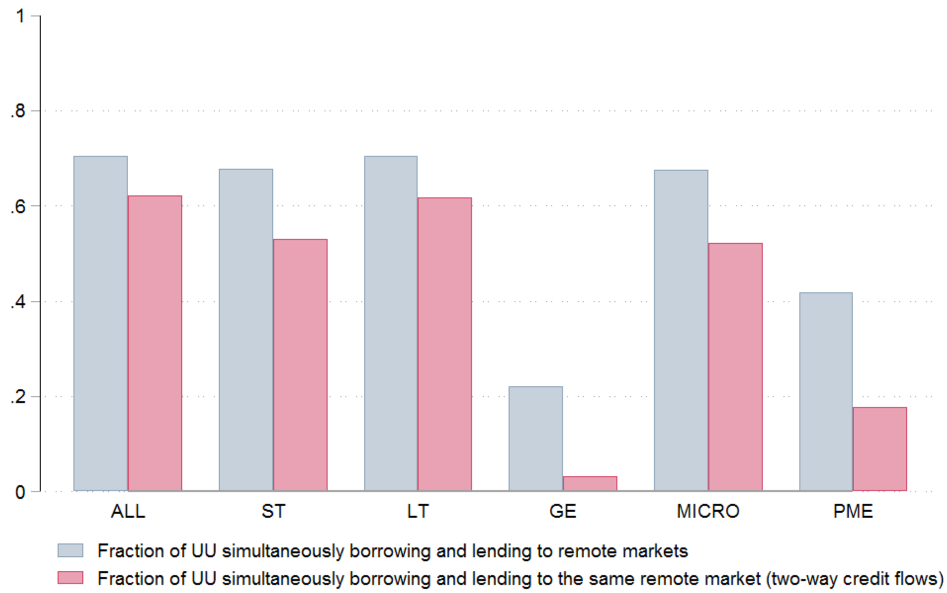
Notes: This figure displays shows the positive correlation between branch size measured as total credit exposure (top panel) and, alternatively, as the number of clients (bottom panel) and average square geographic distance between the branch and its clients, in kilometers, for the last quarter of 2005.

Fig. 1.9. Average Loan Locally and Number of Distant Markets Penetrated



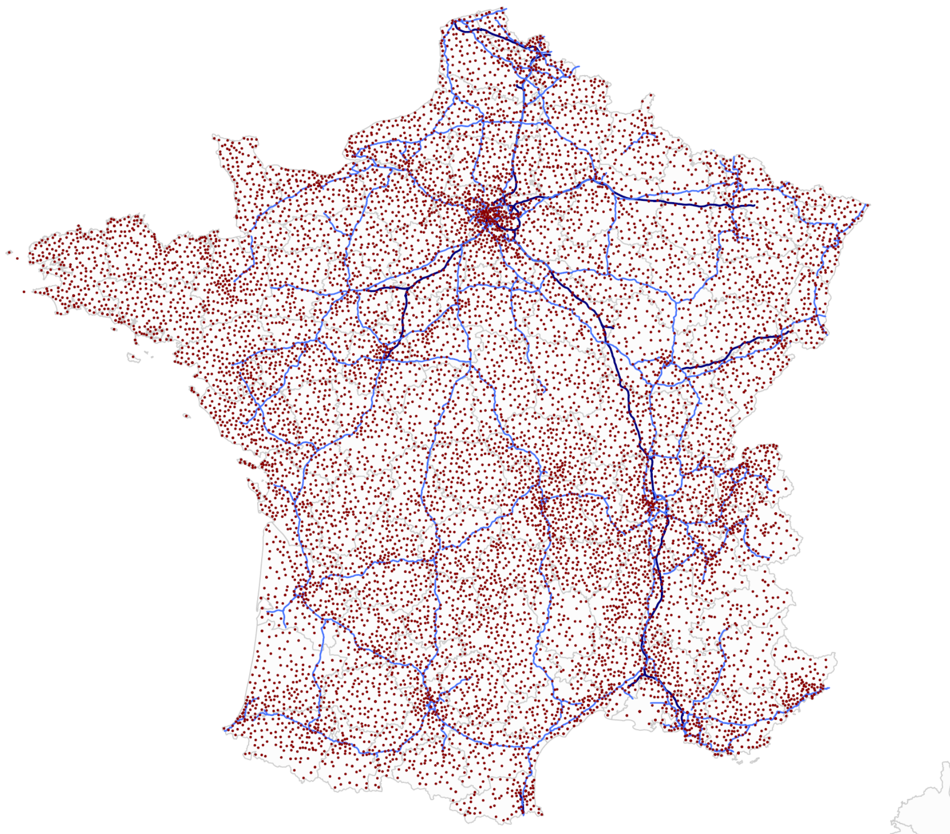
Notes: This figure displays the log average credit size of a branch in its local submarket ("at home") for the group of branches that operate at least in k remote submarkets, with k on the x axis. Bank branches are ranked and grouped based on the number of remote submarkets (i.e. urban units) where they operate: all the branches lend at least to one submarket while none are active in all.

Fig. 1.10. Inter-Submarket and Two-Way Lending



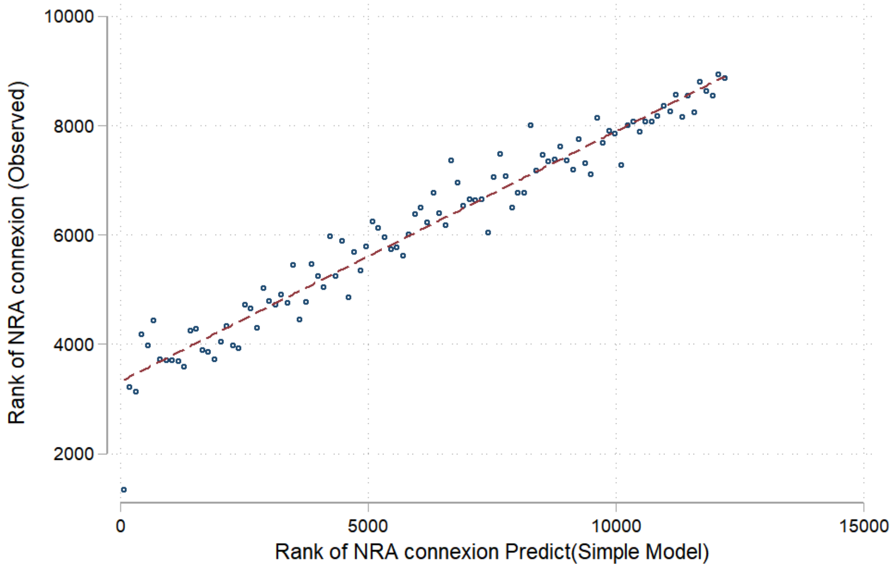
Notes: This figure shows the share of urban units that *borrow* and *lend* simultaneously to remote submarkets (in blue). ALL indicates that all types of credit and all type of clients are included in the computation, while ST stands for short-term credit, LT for long-term. GE indicates that urban units *borrow* and *lend* simultaneously to large firms, MICRO to very small and PME to medium size firms. The red bars indicate the share of urban units simultaneously *borrow* and *lend* to the same distant submarket, with similar sub-categories.

Fig. 1.11. Local Exchanges, Highways and Railroads before 1999



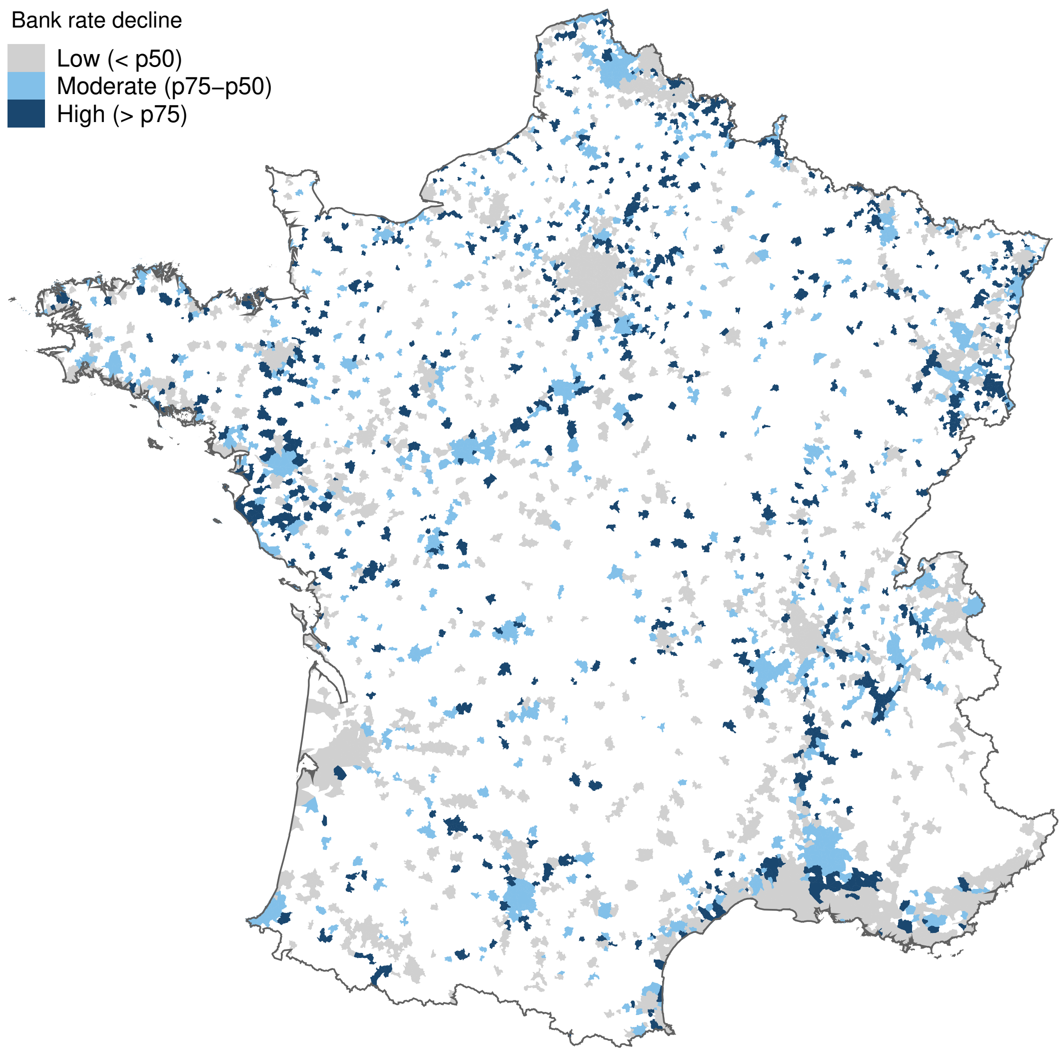
Notes: This figure displays the location of around 13,000 Local Exchanges (red dots), highways (light blue lines) and railroads (dark blue lines) already existing before the beginning of Broadband Internet expansion in France.

Fig. 1.12. Optimal connection rank predicted vs. observed connection rank



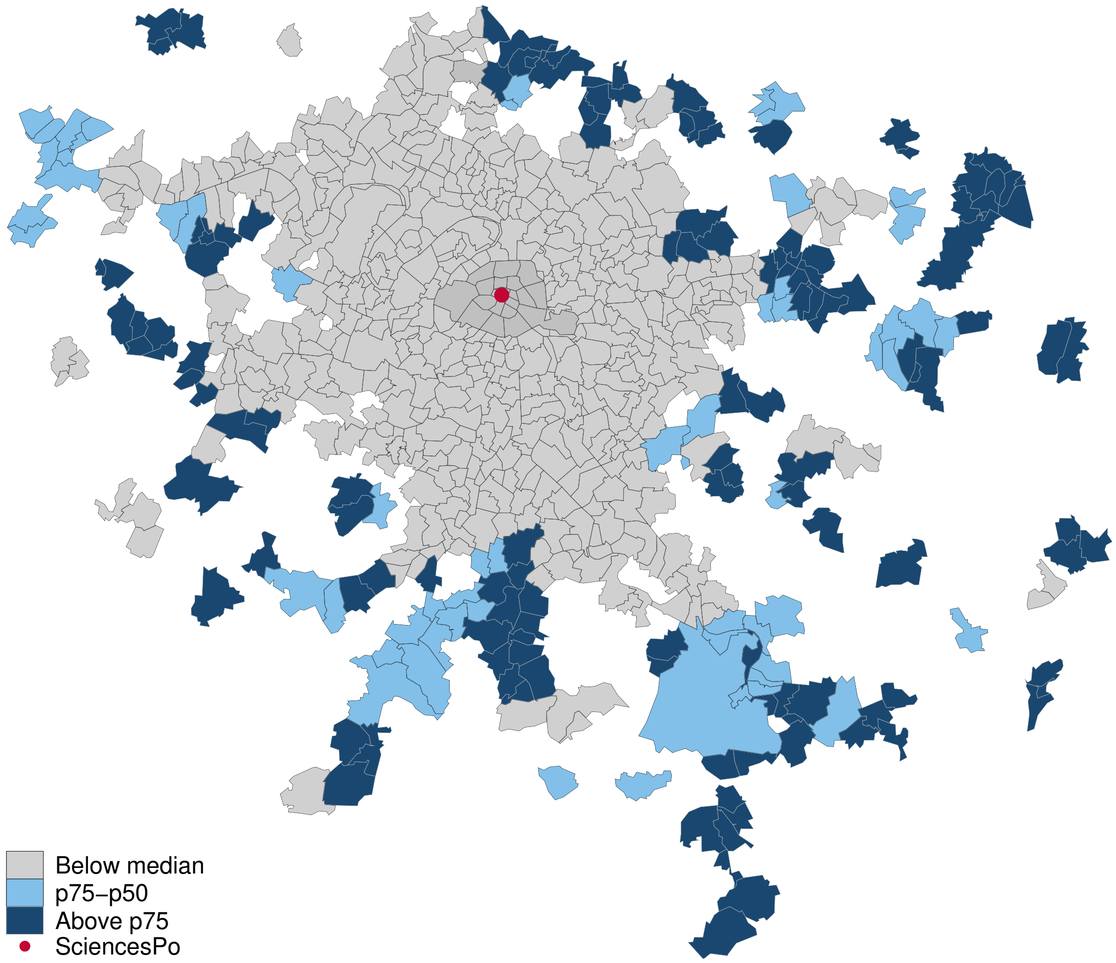
Notes: This figure shows the rank correlation between optimal versus observed connection ranks. The combination of exogenous connection *gains* and *costs* have a strong predictive power, with a R-square close to 0.70. The *optimal connection rank* $\hat{R}_i$ is predicted for each Local Exchange i, only taking into account two presumably exogenous measures of costs (shortest distance to existing infrastructure) and gains (population density).

Fig. 1.13. Reduction of the cost of debt triggered by a reduction in search frictions: spatial heterogeneity



Notes: This map illustrates the spatial dimension of the decline in firms' cost of debt in France. Dark blue cities are the one in which this decline is the strongest, larger than the 75^{th} percentile. The light blue areas undergo a decline of the cost of debt higher than the median (but lower than the 75^{th} percentile). Finally the light grey ones indicates a reduction lower than the median.

Fig. 1.14. Reduction of the cost of debt triggered by a reduction in search frictions: zoom in Paris region



Notes: This map illustrates the spatial dimension of the decline in firms' cost of debt with a focus on the Paris region. Dark blue cities are the one in which this decline is the strongest, larger than the 75^{th} percentile. The light blue areas undergo a decline of the cost of debt higher than the median (but lower than the 75^{th} percentile). Finally the light grey ones indicates a reduction lower than the median. The red dot indicates the localisation of SciencesPo, i.e. the center of Paris.

10. Tables

Table 1.5: Gravity Equation for Inter-regional Credit Flows

	Share of credit in v borrowed from u: $\Pi_{uvt} > 0$				
	(1)	(2)	(3)	(4)	(5)
Log distance d_{uv}	-1.001***	-0.992***	-0.934***	-0.971***	-0.955***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Same region		0.081***			
		(0.013)			
Same county			0.419***		0.669***
			(0.010)		(0.016)
log (Trade Flows)				0.047***	-0.094***
				(0.003)	(0.005)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
Observations	143,747	143,747	143,747	143,747	143,747

Notes: PPML estimation of equation (1.23). d_{uv} = bilateral distance. Same region is a dummy variable equal to 1 if both the firm and the bank are located in the same region. Same county is a dummy variable equal to 1 if both the firm and the bank are located in the same county. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997-2005. The sample consists of origin-destination-year combinations with positive credit flows, where at least one firm is located with positive credit.

Table 1.6: Technology-Induced reduction in search frictions: Pair fixed-effects

	Share of credit in v borrowed from u: Π_{uvt}			
	(1)	(2)	(3)	(4)
Log distance d_{uv}	-1.844***		-1.865***	
	(0.006)		(0.006)	
$\mathbb{C}_{uvt}$	0.800***	-0.009		
	(0.059)	(0.033)		
Log distance d_{uv} x $\mathbb{C}_{uvt}$	0.230***	0.035***		
	(0.008)	(0.004)		
$\hat{\mathbb{C}}_{uvt}$			0.399***	0.037***
			(0.062)	(0.004)
Log distance d_{uv} x $\hat{\mathbb{C}}_{uvt}$			0.199***	0.108***
			(0.008)	(0.039)
log (Trade Flows)	0.470***	-0.018**	0.468***	-0.018**
	(0.004)	(0.009)	(0.004)	(0.009)
Origin (u) $\times$ Year FE	Yes	Yes	Yes	Yes
Destination (v) $\times$ Year FE	Yes	Yes	Yes	Yes
Origin-Destination FE	No	Yes	No	Yes
Observations	24,545,640	250,780	24,545,640	250,780

Notes: PPML estimation of equation (1.24). d_{uv} = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$, is a continuous variable that indicates the degree of broadband internet inter-connectivity between two urban units. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My instrument variable strategy delivers a similar measure of between urban unit connectivity named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) and (3) include fixed effects for origin $\times$ years and destination $\times$ years, columns (2) and (4) add pair fixed-effects. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

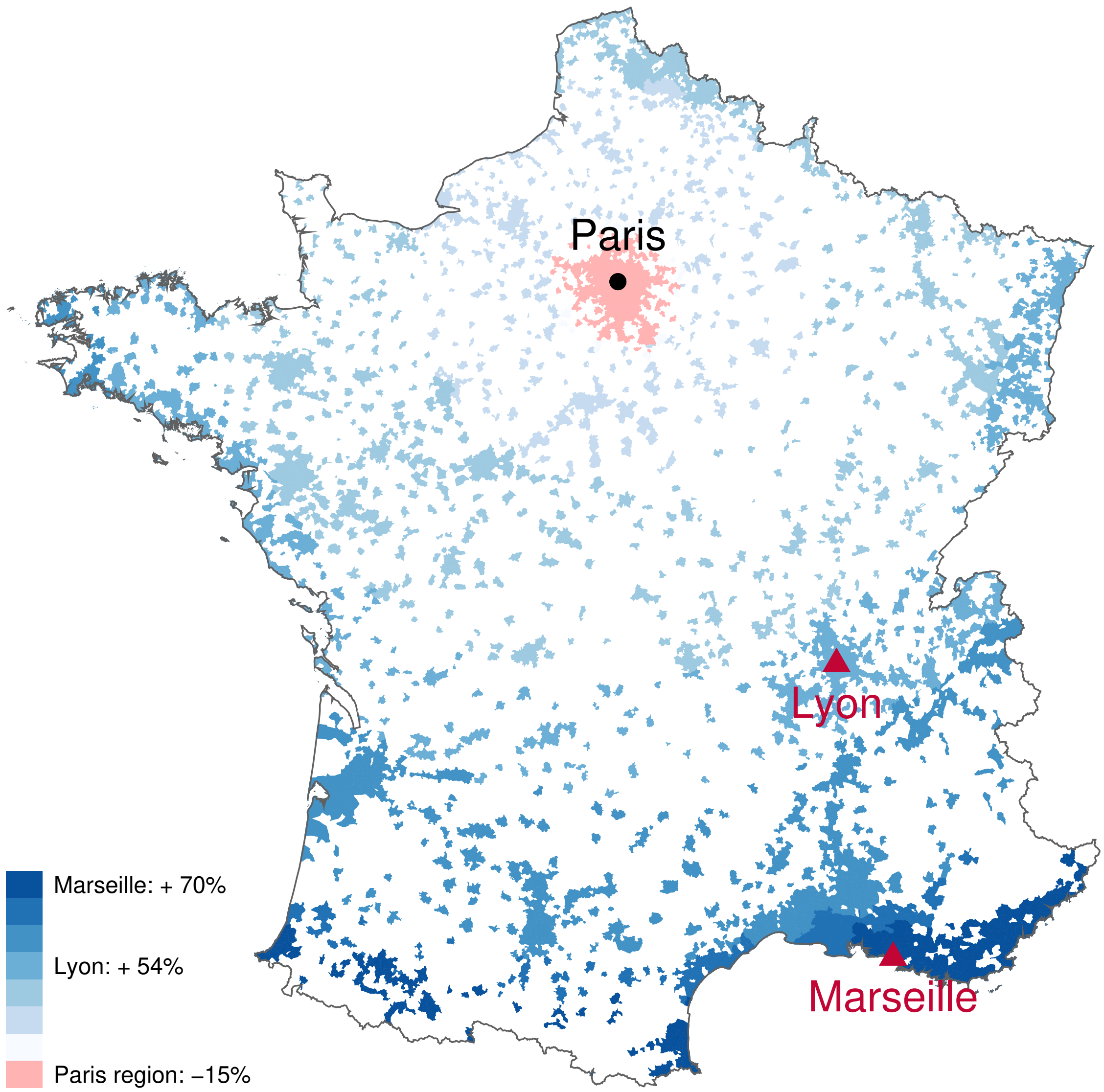
Table 1.7: Technology-Induced reduction in search frictions with lags

	Share of credit in v borrowed from u Π_{uvt}					
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance d_{uv}	-1.561***	-1.764***	-1.175***	-1.265***	-1.174***	-1.274***
	(0.003)	(0.004)	(0.004)	(0.006)	(0.004)	(0.006)
$\mathbb{C}_{uvt}$			0.097*	0.416***		
			(0.053)	(0.053)		
Log distance $d_{uv} \times \mathbb{C}_{uvt}$				0.214***		
				(0.007)		
$\hat{\mathbb{C}}_{uvt}$					0.107**	0.286***
					(0.052)	(0.053)
Log distance $d_{uv} \times \hat{\mathbb{C}}_{uvt}$						0.168***
						(0.007)
log (Trade Flows)		0.473***	0.482***	0.478***	0.482***	0.477***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Origin (u) $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination (v) $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No	No
Lag dependant variable	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24,545,640	24,545,640	20,633,055	20,633,055	20,633,055	20,633,055

Notes: PPML estimation of equation (1.24) with lag dependant variable included. d_{uv} = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$, is a continuous variable that indicates the degree of broadband internet inter-connectivity between two urban units. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My instrument variable strategy delivers a similar measure of between urban unit connectivity named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin $\times$ years and destination $\times$ years. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

A. Appendices

Fig. 1.15. Heterogeneous effect of BI with respect to distance when Paris is connected



Notes: This figure plots the effect of broadband internet connection on the share of credit borrowed by firms located in any city in France to banks located in Paris. The black dot indicates the Paris location, while the red triangles shows Marseille and Lyon, respectively the second and the third biggest french cities. Dark blue indicates an effect in the the 90th percentile while light red indicates the negative effect of being connected to Paris for cities located close to Paris.

Table 1.8: PPML with many zeros in a dynamic setting: simulation results

Data Generating Process I Obs. = 200 x 200 x 10, a = 50 b = 0
$\mathbb{P}(Y_{uvt} = 0) = 0.97$

Estimator	β_1		β_2		β_3	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PPML	-1.003	0.029	1.001	0.037	0.226	0.033
PPML w. D×T + O×T FE	-1.006	0.029	1.006	0.084	0.121	0.099
GPML	-0.874	0.034	1.008	0.039	0.185	0.035
GPML w. D×T + O×T FE	-1.019	0.03	0.995	0.086	0.093	0.101

Data Generating Process II Obs. = 200 x 200 x 10, a = 1 b = 5
$\mathbb{P}(Y_{uvt} = 0) = 0.88$

Estimator	β_1		β_2		β_3	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PPML	-1.001	0.013	1.001	0.017	0.226	0.015
PPML w. D×T + O×T FE	-1.001	0.013	0.999	0.038	0.115	0.045
GPML	-0.872	0.018	0.998	0.019	0.187	0.017
GPML w. D×T + O×T FE	-1.005	0.015	0.996	0.041	0.11	0.04

Notes: This table presents the results obtained with 1,000 replicas of the simulation procedure described in Section 5. The number of cities N is set to 200 and T equals 10. The estimation sample is therefore composed of 200 × 200 × 10 = 400,000 observations. In the top panel, a = 50 and b = 0 while a is set to 1 and b to 5 in the bottom panel. The table displays the average point estimate and the standard errors obtained with different estimators, namely PPML and GPML with and without origin × year and destination × year fixed-effects. Robust standard errors.

2

Bank Local Specialization

This chapter is based on a paper co-authored with Anne Duquerroy (BdF), Jean-Stéphane Mésonnier (BdF) and Daniel Paravisini (LSE, CEPR).

Abstract

Using micro-data of the universe of bank-SME relationships in France, we show that banks specialize locally (at the branch level) by industry, and that this specialization shapes the equilibrium amount of borrowing by small firms. For identification, we exploit the reallocation of local clients from closed down branches to nearby branches of the same bank, which induced quasi-random variation in the match between a firm's industry and the industry of specialization of the lending branch. We show that branch reallocation leads, on average, to a substantial and permanent decline in small firm borrowing. This decline is twice larger for firms whose accounts are reallocated from branches less specialized in their industry than the original one.

JEL classification: G21

Keywords: bank specialization, SMEs, relationship banking, branch closures.

Introduction

Widespread bank branch closures and consolidation in Europe and the United States after the Great Recession have renewed a longstanding policy and academic debate about the nature and implications of bank market power. A vast body of theoretical and empirical work, following Rajan (1992), has explored one source of such market power: the informational monopoly gained through relationship lending. Less studied, and of potentially equal importance, is the market power gained through differentiation and specialization. Lenders that may appear to compete fiercely in an undifferentiated credit market, may in fact enjoy market power in some market segments by tailoring their products and services to particular clients, industries, or types of financing. Such credit

market segmentation may have first order implications for the access to credit by small and opaque bank-dependent firms as well as its cost. Documenting the extent to which banks specialize in a segmented small business credit market, and assessing whether specialization confers market power, poses important data and identification challenges that we address in this paper.

We use unique regulatory data that contains, for the universe of bank-firm relationships in France, the identity and location of the bank branch providing credit. With these data we construct measures of bank and bank-branch industry specialization in narrowly defined geographical credit markets. Figure 2.1 provides stylized motivating evidence for our study. It plots two different measures of credit market concentration by urban unit.[1] The left panel shows the standard concentration measure, calculated using total lending shares. The right panel shows the average of credit concentrations calculated industry-by-industry, which takes into account market segmentation. The difference between these two measures will be larger when the credit market is segmented by industry (e.g., the two measures will be identical if all bank loan portfolios have the same industry composition). Indeed, the fraction of urban units with a very high concentration level (HHI > 0.4) raises from from 21% when measured in the traditional manner, to 49% when measured taking into account credit segmentation by industry (see Table 2.9 for more details). This pattern is consistent with heavy industry segmentation in the small business bank credit market.

The main goal of our empirical analysis is to explore the implications of credit market segmentation for small firms' access to credit. Our working hypothesis is that a firm's credit elasticity of substitution across banks is smaller when specialized banks offer differentiated services. For example, a firm in the construction industry will find more difficult or costly to substitute credit obtained from the bank (or branch) that is specialized in the construction industry than credit obtained from a generalist bank (or branch). A necessary first step is to evaluate the relevant unit of analysis to study specialization. In other words, do banks specialize by industry as a whole? Or is specialization a local, branch-level, phenomenon?

To answer this question we follow the data driven approach developed in Paravisini et al. (2017) to identify banks' sector of specialization using abnormally large portfolio shares. The intuition of the measure is best explained through an example. Suppose 20% of bank credit in an urban area goes to the construction industry and is serviced by five banks. Banks are heterogeneous: while four banks allocate less than 10% of their loan portfolio to construction, the fifth bank allocates more than 40% of its credit portfolio to the sector. This fifth bank

[1] An urban unit is defined as a municipality or a group of municipalities which covers a continuously built up zone (with no more than 200 meters between two constructions) and hosts at least 2,000 inhabitants.

would be identified as a specialist in the construction industry for this urban area. The advantage of using portfolio shares to detect specialist banks is that the identification of the specialization sector is unaffected by the size of the sector or by the market share of each bank in any given location.

Two key stylized facts emerge from this exercise: bank branches tend to specialize by industry, but different branches of the same bank generally exhibit different industrial specializations. More than a third of bank branches in France come out as being specialized in supplying credit to small firms in at least one specific broad industry. Most urban areas include specialized bank branches. Moreover, we observe that most industries exhibit specialized bank branches at the local level. For instance, some 9% of the bank branches present in our sample in 2017 are specialized in funding transportation and storage activities. Overall, this implies that a French SME has a non-negligible probability to get connected to a branch that is specialized in its type of business. When we investigate specialization patterns of the branches within banks, we find that large banks are characterized by a large share of specialized branches (37% for the average bank with more than 10 branches). However, within a bank, different branches tend to be specialized in different industries. In short, industry specialization appears in the data as a widespread but local, branch-level, phenomenon.

Motivated by these stylized facts we turn to measuring the heterogeneity in firm's elasticity of credit substitution by branch specialization. Our empirical research design exploits borrower reallocations across branches due to branch closures. Among bank branches active in SME lending, some 700 branches were closed during our sample period (between 2010 and 2017) throughout the country, due to internal restructuring plans of the main banks' retail activities. Branch closures did not end bank-borrower relationships: all loan accounts in a closing branch were transferred to larger nearby branches of the same bank. Branch reallocation induced variation in the match between the borrower's industry and the industry of specialization of the branch that we exploit to measure the heterogeneity in the elasticity of credit substitution. In the construction firm example above, when branch services are segmented by industry, the transfer of the firm's account to a generalist branch should reduce the equilibrium amount of credit used by the firm, relative to a counterfactual in which the account were transferred to another branch that is also specialized in the construction industry. Branch closures occurred in large waves, and the identity of the closing and absorbing branches were selected by headquarters according to criteria like local bank density, arguably unrelated to the demand for credit of individual firms. The very disaggregated nature of the data also allows using saturated specifications to control for local shocks at the urban unit level, bank shocks, and firm shocks that may occur concurrently with the branch closure.

In the baseline specification we find evidence of a significant drop in the total of credit granted by a bank to a small firm whenever the firm's account is reallocated to a new branch. Including undrawn credit lines, total credit drops by 12% on average over the three years following the effective closing. Part of this decline is substituted with more credit from other banks. However, the average firm's total credit drops permanently by about 4% after an account reallocation, relative to other firms in the same narrow geographical market and industry. We then document the heterogeneity of this decline in equilibrium credit by the match between the borrower's industry and the industry of specialization of the closing and absorbing branches. We find that the magnitude of the decline in credit doubles when a firm's accounts are reallocated from a branch that specializes in its industry to a branch that does not. The magnitude of this estimated effect is robust to controlling for the change in distance associated with the branch closing. In the cross section, we find that the decline in credit following a branch closure is entirely explained by the variation in industry specialization between branches when the new branch is located in an area characterized by a high level of bank competition. The results are strongly suggestive of a segmented bank credit market, where bank specialization by industry increases the cost of substituting bank sources of financing for small firms.

Related Literature. The results in this paper complement an extensive literature on relationship lending to small firms (for surveys see, e.g., Boot (2000); Ongena and Smith (2000); Degryse et al. (2009)). This literature analyzes the market power banks gain thanks to the private information gathered through the lending interaction with small and opaque borrowers.[2] We extend this literature by characterizing the complementary role of the industry segmentation of the small-firm credit market.

The results represent a novel contribution to the nascent literature studying the extent and consequences of bank specialization in corporate lending. Studies that incorporate specialization and segmentation in the analysis of bank competition in the market for corporate credit is scarce, a stark contrast to its widespread adoption in the study of consumer banking services, such as mortgages (e.g., Benetton (2020)), pensions (e.g., Hastings et al. (2017)), and deposits (e.g., Egan et al. (2017)), amongts others. Of the existing work on bank specialization, the closest to ours is Paravisini et al. (2017), which shows that lenders specialize by export destination market in the context of Peruvian exporters.[3] Our contribution to

[2] See, for example, Sharpe (1990); Rajan (1992); Berger and Udell (1995, 2002); Agarwal and Hauswald (2010); Ioannidou and Ongena (2010); Degryse and Ongena (2005); Drexler and Schoar (2014); Petersen and Rajan (1994); Nguyen (2019))

[3] Two other recent studies that incidentally point to a potential role for banks' industry specialization are Ongena and Yu (2017), which finds that multi-industry firms tend to borrow from a larger number of banks, and that they tend to pick up banks according to their industry specialization, and De Jonghe et al. (2019),

this work is twofold. First, we show that the effect on competition of market segmentation through differentiation is of first order magnitude for small opaque firms. And second, we uncover and document the local and decentralized nature of specialization within large banks.

These findings combined imply that the true extent of market power, segmentation, and specialization in the bank credit market for small firms can be obscured by aggregate or bank-level data. Also, the results have important implications for the consequences of bank branch consolidation, which motivates this paper. The mechanism uncovered here is related to, but economically distinct from, the role played by physical distance and emphasized by existing work (e.g., Nguyen (2019)). In very related work, Bonfim et al. (2016) documents that loan conditions change when firms switch to a new bank following to closure. Our results highlight that branch closures affect the credit relationship even when borrowers remain with the same lender, through changes in local competition.

The remainder of this paper is organized as follows. Section 1 describes the spread of French banking networks and institutional details about branch closures. Section 2 describes the data used in our analysis. In Section 3, we present our measure of bank branch specialization and provide descriptive statistics of this new variable. Section 4 describes our empirical strategy and section 5 details the results of our analysis. Last, Section 6 concludes.

1. Bank branch networks in France

France has one of the most developed network of bank branches in the euro area. Bank branches are geographically widespread throughout the country, leaving almost no room for so-called banking deserts, even though the largest urban areas concentrate a larger number of bank branches. As an illustration, the left panel of Figure 2.2 shows the spatial distribution of branches lending to SMEs as of December 2016. However, in the past couple of years, major French banks have announced plans to reduce their branch network by 2020 as customers are walking into their bank less often than before. These plans confirm and accelerate a trend that started with the financial crisis in the previous decade, against the backdrop of a concomitant rise of online banking. In aggregate, the overall number of bank branches, lending to both households and firms in France, has declined by 4%, from 38,784 branches

which finds that Belgian banks hit by the Lehman shock tend to shore up borrowers in industries where the bank is dominant and industries where it is specialized. In addition, Goedde-Menke and Ingermann (2020) find that lower specialization - defined as a higher diversification of loan officers' portfolio - following a downsizing in the bank staff of a German bank, translates into higher risk taking in the intensive margin.

in 2010 to 37,261 in December 2016 [4].

Among these branches, some only lend to corporate clients, some only serve households and others serve both. In this paper, we focus on branches that distribute credit to small and medium-size enterprises (SMEs).[5] At the end of 2016, about a third of all bank branches (12,291 branches) were actively lending to SMEs [6]. Among the branch closings registered since 2010, some two thirds happened as the consequence of a merger between two banks. We focus here on the remaining closings, that were not related to bank mergers, so that the post-closing transfer of firm accounts remains always confined within the same initial relationship bank. From 2010 Q1 to 2017 Q3, some 700 branches that were active in local SME lending closed in continental France as a result of internal bank strategies.[7] These closings have been widespread over time and geographically dispersed as illustrated by figures 2.3 and 2.4. Figure 2.3 shows the spatial distribution of all branch closings over the 2010-2017 period. Over our sample period 45% of branch closings took place in small cities (urban units with less than 20,000 inhabitants) and 15% in large cities (more than 100,000 inhabitants).

In practice, when a bank decides to close a branch, the closing announcement takes place 14 to 17 months before the branch officially closes and actually stops operating. Following the closing of the branch where their account is managed, borrowers do not lose access to their bank: they are automatically transferred to another branch within the same bank. They do not get to choose their new branch in the vast majority of cases. On average 90% of firms in our sample get transferred to the absorbing branch while the rest of them may have made their own decision to switch to another entity.

In our analysis of the effects of local bank specialization, we exploit branch closings as events which trigger exogenous variation in the industry specialization of the bank branches to which a given firm is connected. For our identification strategy to be valid, it is therefore important to ensure that branch closings are not on average a consequence of a lower credit quality of the branch's local borrowers. If this were the case, any reduction in credit received by a firm

[4]Source: ECB structural banking indicators.

[5]We define firm size according to the European definition. SMEs are enterprises with less than 250 workers and a turnover or total assets not exceeding, respectively, 50 million euros and 43 million euros.

[6]This number is based on lending to SMEs - excluding micro-firms and individual entrepreneurs - recorded in the French Central Credit Register (see data section for more details).

[7]Information about closed branches is consistently registered over time for most French banks but not for all of them. The bank needs to follow a permanent account number policy (so called *RIB*) for panel data information to be available about the demography of its branches. The historical information about the extent of a bank's network that is stored by the Banque de France is limited to 14 months unless the bank has opted for a specific account registration framework called "RIB permanent" (permanent bank account identifier), henceforth RIB-p banks. Our sample of branch closings events is thus restricted to such banks for the period 2010- 2014 and exhaustive starting in 2015 as we have started to collect and record information on branch closings on a regular basis at the beginning of 2016.

following the transfer to a new branch could reveal its lower quality, which would eventually be uncovered by the loan officers of the new branch, rather that a drop in credit supply by the bank. We therefore need to ascertain that closures are not merely the consequences of adverse local economic conditions weighing down on local borrowers.

Unfortunately, banks do not disclose the precise criteria they apply when deciding which branches have to close. When they communicate upon the topic, e.g., in their annual reports, banks generally motivate the downsizing of their branch network by general developments such as the rise of online banking, and do not refer to local conditions. To gather more direct evidence that closing decisions are independent enough of local economic conditions, we therefore conduct a few preliminary tests.

We first run a standard analysis of variance, whereby we regress a dummy for branch closures (over the period 2010-2017) on various sets of fixed effects. Spatial dummies standing for counties (in French: *départements*) and (much smaller) urban units explain respectively 3% and 8% of the variance of the branch closing dummy.[8] Banking group fixed effects also acount for only a tiny share of variance. In contrast, the bank-level fixed effects explain alone 32%. Overall, this suggests that the decision to close a branch depends more on banks' global strategy rather than on local factors.

Second, we estimate a linear probability model that relates the probability that a branch closes to a set of (i) branch-level variables and (ii) measures of local economic dynamics. The dependent variable is again a dummy equals to 1 if the branch is closed during the 2010-2017 period. Explanatory variables, measured before the period of interest, are the size of the branch (total amount of corporate loans, in log), the branch's market share in the county, a dummy for being the only local branch of the bank, the rank of the branch within the bank (of within the bank $\times$ county) and the county's size and population dynamics (measured as the increase in the number of inhabitants over the period 1990-2006).[9] We find that the probability of a closing decreases when the branch is the sole branch of its bank in the county and increases when the county is more populated and its population is on the rise. These results are consistent with anecdotal evidence in the press that banks close more branches in dynamic cities where the population is younger and keener on using online banking. Branch characteristics, as the branch's market size or its relative size within the bank (its rank), seem not to impinge on the decision to close it. Interestingly, the significance of branch-level variables vanishes whenever we add bank fixed effects in the regression. The low predictive power of these branch-specific variables conditionally on bank fixed-effects provide support

[8]Detailed results are shown in Table 2.10 in the online appendix.
[9]Detailed results are shown in Table 2.11 in the online appendix.

to our identification strategy.

2. Data

2.1. Data sources

We merge three different types of information to conduct our empirical analysis: bank-firm credit exposures, bank branch opening and closing dates and geographical data on French municipalities, urban units and urban areas.

First, we get virtually all credit links between domestic bank branches and non-financial firms in France from the French credit register, which records bilateral credit exposures at the branch-firm level above a small reporting threshold of EUR 25,000.[10] A firm is defined here as a registered legal entity. Each firm is identified with a unique legal identifier (SIREN code). A credit exposure towards a borrower is defined as the total of loans, undisbursed credit lines and guarantees given by the bank branch.[11] Some detail is also reported about the outstanding amounts of specific types of loans (notably, leasing, factoring and securitized loans). This information is collected at monthly frequency. To limit the size of our database, we however keep end-of-quarter observations only. Banks are individual credit institutions that have been authorized to operate in France (including branches of foreign banks). They are identified by a 5-digit code (BIC code) and can be part of a banking group. We map banks into their respective consolidating groups (*Groupes economiques d'appartenance*, or GEA) using additional information from the French supervisor. The credit register also includes the ZIP codes of both the borrowing firms and the bank branches, which allows us to compute geopgraphical distances between borrowers and lenders.

Second, data on branch closures are obtained from the Banque de France FIB database (*Fichier des Implantations Bancaires*) since 2010. This database collects information on every single bank branch operating in France, including its precise location (ZIP code/street address), its opening date and the BIC number of the parent bank. As regards branch closures, the database records the official date of the closing, the date of the closing announcement, as well as the reason why a branch was closed (merger or acquisition by another bank

[10] The threshold applied at the branch-firm level up to March 2012, and at the (less restrictive) bank-firm level thereafter. We imposed this threshold at the branch-firm level throughout the 2010-2017 period to avoid a statistical break.

[11] Two banks, *La Banque Postale* (the French post bank) and another French cooperative bank, only report their bilateral credit exposures at the county (*departement*) level, and not at the level of local branches. We drop these two banks from our sample. They accounted for less than 1% of outstanding amounts of loans to SMEs in December 2016.

or internal bank reorganization). In the case of a within-bank closing and a subsequent transfer of accounts, the identity of the absorbing branch, as well as its location, are reported. For our purpose, a firm is affected by the closing of bank branch (treatment) if this firm had a lending relationship with the closing branch in the 12-month period before the official closing date. We drop firms affected by more than one closing event (less than 10% of treated firms).

Last, we gather geographical data for mainland France about some 36.000 municipalities (identified with a ZIP code), 2,000 urban units and 771 urban areas. A urban unit (UU) is a municipality (*commune*) or a group of municipalities that includes on its territory a built-up area of at least 2,000 inhabitants where no dwelling is separated from the nearest one by more than 200 meters. In addition, each municipality has more than half of its population living in this built-up area. The largest geographical unit that we consider in this paper is the urban area (UA). A urban area is defined as a group of municipalities, all in one piece and without enclaves, consisting of a urban pole with more than 10,000 jobs, and rural municipalities or urban units (peri-urban crown) where at least 40% of the resident population with a job works in the pole or in municipalities attracted by it.[12] We compute firm-branch distances with reference to the centroid of the municipalities they are located in. In contrast, we use urban units and urban areas in order to control for geographical dynamics at a broader and more economically relevant level in our empirical analysis.

2.2. Data cleaning

We aim to analyze credit relationships between small and medium-sized businesses (SMEs) and the local branches of their relationship banks. The French credit register contains a lot of information that is not relevant for our purpose and it therefore deserves some cleaning.

As far as banks are concerned, we drop all banks for which leasing activities represent more than 95% of their total exposure over the decade 2007-2017. Leasing is a non-standard lending activity that requires specific knowledge and expertise. This specific financial service is most often provided by specialized institutions (in general subsidiaries of larger banking groups) that operate remotely, mostly from Paris. Even though the contract with the leasing company may be sold by the local branch of a deposit-taking bank that belongs to the same group, the issuing branch is recorded in the credit register as being the branch of the leasing institution and not the local bank branch that only acts here as an intermediary. As we aim to study local bank-firm relationships, these leasing institutions do not enter into the scope

[12]The urban zoning also distinguishes "medium areas", a group of municipalities consisting of a pole with 5,000 to 10,000 jobs and "small areas", a group of municipalities consisting of a pole with 1,500 to 5,000 jobs

of this analysis. Additionally, we drop all French public financial institutions (e.g. Groupe CDC, Caisse nationale des autoroutes, BPIFrance etc.). Last, we exclude bank branches located in Corsica and in the French overseas territories (*DOM-COM*). Individual banks are mapped into their respective banking groups (so-called *Groupes Economies d'Appartenance*) using additional information provided by the French supervisor (ACPR).

As far as borrowers are concerned, we first exclude sole-entrepreneurships so as to focus on corporations. We drop firms that belong to the financial sector, local public administrations, non-resident firms, as well as real estate companies (*SCI*).[13] We consider in our final sample only legal entities belonging to small and medium-sized corporations (SMEs) as defined by the French LME Act of 2008, i.e. firms with less than 250 employees and a turnover below EUR 50 mns (or total assets below 43 mns).[14] However, we exclude micro-enterprises (firms with less than 10 employees and an annual turnover below EUR 2 mns) in order to keep the size of our sample manageable. We also drop firms with missing size classification. We focus on firms with at least two banking relationships. Last, we only keep firms that are located in urban units where at least one branch closing occurs and are present at least 8 quarters in a row (2 years) in the credit register.[15] We further impose that treated firms borrow funds from the closing bank branch during at least one quarter in the year preceding the *announcement* of the closure (quarter "-11" to quarter "-7").

2.3. Descriptive statistics

We conduct our baseline empirical analysis using credit exposure data aggregated at the firm-bank level, while keeping a record of bank branch closures that affect each bilateral relationship in order to construct our treatment variable. Table 2.1 provides detailed statistics of our firm-bank dataset. Our estimating sample consists in more than 5.2 millions firm-bank relationships, over the period 2010 Q1 to 2017 Q3. We observe on average 346 banks, corresponding to 11,786 bank branches, and 77,640 firms per quarter. The average number of firm-bank relationships in the cross-section is close to 190,000. The average credit relationship involves an outstanding loan amount of some EUR 430,000 (some EUR 530,000 when undrawn credit lines are taken into account), but half of the bilateral bank-firm relationships involve much smaller amounts, below EUR 115,000. The table also confirms that

[13] We likewise delete observations for various legal categories under the French civil, commercial or administrative law that are irrelevant for our analysis (eg. parishes, unions, some types of cooperatives, etc.).

[14] The legal definition of an enterprise is the smallest grouping of legal entities that makes up a coherent production unit of goods or services and enjoys some minimal degree of management autonomy.

[15] This means that the total of all loans (including short and long term credit, credit-lines, overdrafts, account receivables, export credit, and leasing and factoring) and bank guarantees received by the firm from all selected banks is non-zero during a period of at least to years.

SMEs tend to pick up lenders in a relatively small neighborhood: the median of the distance between a SME and its bank in our dataset is only 14 km.[16] In Table 2.2, we also show firm-level statistics for our sample of multi-bank SMEs. The outstanding bank debt of the average firm is EUR 1 million, which confirms that most of the firms in our sample are rather small. The average total amount of credit granted to a firm, including undrawn credit lines, is slightly larger (EUR 1.2 million) and the average ratio of long-term credit to drawn credit is 50% (where long-term credit refers to credit with initial maturity over one year).

3. Measuring local bank specialization

3.1. Measuring local bank specialization

In this section, we detail how we define and compute local bank specialization and provide some descriptive statistics of this new variable. We are interested in this study in the industrial specialization of bank branches. Our definition of specialization relates to Paravisini et al. (2017), although we define here specialization at the level of bank branches instead of banks and we consider specialization vis-à-vis local borrowers in a specific industry instead of exporters to a given country. Note that, in order to assess the local industry specialization of bank branches, we use information about all bank branches lending to non-financial firms in continental France between 2010 and 2017 and measure their lending activity towards all domestic non-financial firms (including large firms).[17]

We proceed in four steps. We first define the credit concentration of branch h of bank b in an industrial sector s as the ratio of credit supplied by branch h to firms in s over the total amount of credit supplied by branch h to firms in all sectors:

$$\text{Concentration}_{b,h,s} = \frac{\sum_{i \in (b,h,s)} L_i}{\sum_s' \sum_{i \in (b,h,s')} L_i} \tag{2.1}$$

where L_i denotes the amount of credit lent by h to a firm i.

Second, the credit concentration of h in s needs to be compared with a relevant local bench-

[16]The distance to the bank is computed as the unweighted average of the distances to all the banks' branches the firm borrows from.

[17]This information set is therefore larger than the sample of bank branches and firms used in the regressions. in practice, we restrict the population to branches located in urban areas with more than 10 branches -which is almost always the case- and branches with more than 5 customer firms. Increasing this threshold to a minimum of 10 borrowers slightly increases the number of branches for which we can compute a specialization indicator but does not affect our results.

mark. For this purpose, we aggregate the loan portfolios of all bank branches located in the same urban *area* u as h. We then define the benchmark concentration of loans to sector s in the neighborhood of h as the share of s in the credit supplied by bank branches located in u (including h). The relative local credit concentration of branch h in sector s then reads:

$$\hat{S}_{b,h,s,u} = \text{Concentration}_{b,h,s} - \text{Concentration}_{u,s} \tag{2.2}$$

This measure lies in the]-1;1[interval. A relative credit concentration in s equal to zero means that share of s in h's credit portfolio is aligned with the composition of lending in the neighborhood, possibly reflecting either a strong presence or the absence of s-type firms in this neighborhood. In contrast, a value close to 1 indicates that branch h almost exclusively funds industry s, while other banks in the neighborhood do not. In that sense, h is specialized locally in financing industry s.

Third, we follow on Paravisini et al. (2017) and focus on outlier values of $\hat{S}_{b,h,s,u}$. More precisely, a branch is identified as being specialized in industry s (i.e., the specialization dummy $S_{b,h,s,u}$ is set to one) if $\hat{S}_{b,h,s,u}$ is above the upper extreme value, defined by the 75-th percentile plus 1.5 inter-quartile ranges, of the distribution of $\hat{S}_{b,h,s,u}$ across all bank branches operating in continental France in quarter q:

$$S_{b,h,s,u} = \mathbb{I}(\hat{S}_{b,h,s,u} \geq \text{p75} + 1,5 \cdot \text{IQR}(F(\hat{S}_{b,h,s,u}))) \tag{2.3}$$

where F($\hat{S}_{b,h,s,u}$) is the distribution of relative lending intensity of all branches located in urban area u. As highlighted in Paravisini et al. (2017), identifying outliers using percentiles and inter-quartile ranges has the advantage that it does not rely on any assumptions about the distribution of bank portfolio shares.

Fourth, when a branch h is closed and all its clients are reallocated to branch h', these firms face an exogenous change in the industrial specialization of their relationship lender. For each firm i, this matters only as far as either h or h' are specialized in the industry of i. We then define a firm-specific *relative loss* in the industry specialization due to a branch reallocation:

$$\Delta S_{b,i,u} = S_{b,h,s(i),u} - S_{b,h',s(i),u}$$

If both branches share the same level of specialization in the industry to which i belongs,

denoted $s(i)$, then $\Delta S_{b,i,u}$ equals zero and i is not affected. If the closing branch is specialized in $s(i)$ and the absorbing one is not, then $\Delta S_{b,i,u}$ equals one and firm i is likely to face higher informational frictions in its access to credit because the new loans officers in branch h' are less skilled in monitoring firms in this specific industry.

In practice, we define industries at the one-digit level of the NACE rev2 classification (refered to using letters from A to S). We therefore measure bank branch specialization in terms of 15 broadly defined sectors, such as agriculture, forestry and fishing (A), mining (B), manufacturing (C), construction (F), wholesale and retail trade and repair of motor vehicles and motorcycles (G), accommodation and food service activities (I) or administrative and support service activities (N). We smooth out some noisy variations due to the small size of some branches and neighborhoods by computing $\Delta S_{b,h,h',s,u}$ as the average specialization of the closing branch the year before the closure minus the specialization of the new branch one quarter before the closure.[18]

3.2. Stylized facts about local bank specialization

In this section, we provide a first description of the industrial specialization of bank branches in France in the 2010s. First, we find that the local industrial specialization of bank branches is a rather common phenomenon. Some 36% of bank branches lending to firms are specialized in lending to at least one industry at some point in time. Moreover, specialized bank branches are not clustered in some region but roughly equally spread across the whole country, as the second panel of Figure 2.2 shows. Specialized branches are present in some 90% of urban units where bank branches lending to firms are located, although they are more numerous in larger cities hosting a larger number of branches. On average over 2010-2017, the probability that a firm is borrowing from a bank branch that is specialized in financing its industry is close to 30%.

Second, we observe that most (broadly defined) industries benefit from the presence of locally specialized bank branches. Figure 2.5 shows the number of specialized bank branches per industry as of September 2017, to be compared to the some 12,000 bank branches in our sample at this date. Some 900 branches come out as being specialized in lending to the transportation and storage sector [H]. About 750 and 700 branches are specialized in lending to firms which operate in the information and communication sector [J] and, respectively, in the construction sector [F]. In contrast, almost no branches are specialized in funding firms selling other services activities [S] or manufacturing firms [C].

[18] The lags also warrant that branch-level specialization measures are not affected by the closure.

The upper panel of Table 2.3 provides a view on branch specialization from the perspective of individual banks as of September 2017. The 294 banks in our sample have between one and 1,152 branches in mainland France. On average, a bank has a network of 35 branches, a dozen of which are specialized. However, these distributions are skewed by the large number of small banks, as the median bank only has one branch in this sample. To get a more informative picture of the situation of so-called "network" banks, which run retail banking operations through a network of local branches and make up the bulk of credit supply to firms, the lower panel of the same table shows the same statistics when we focus on the sub-sample of the 81 banks with at least 10 branches. The average network bank has some 120 branches lending to non-financial firms, while 37% of these branches are specialized in supplying credit to at least one industry.

Finally, we ask whether specialized branches within a bank tend to be all specialized in the same industry, which would suggest that the coutrywide specialization pattern of the banks prevails over what we would mistakenly view as local specialization patterns. To get a sense of this, we first compute within each urban unit the share of specialized bank branches which do *not* share the same industry specialization locally as their parent bank coutrywide (e.g., a branch specialized in funding transportation firms which belongs to a bank specialized in funding manufacturing industries fits in this category). Figure 2.6 shows convincingly that local bank branch specialization is largely disconnected from the whole bank's specialization. In the most populated areas, such as the largest urban centers or the French Riviera, the majority of specialized branches are indeed not aligned with the specialization of their parent bank.

Second, we compute for each bank the Hirschman-Herfindahl index of its number of specialized branches in each of the fifteen industries. Conceptually, the index takes the value of one whenever all the bank's specialized branches are specialized in the same industry. Conversely, the index takes a value close to zero (the inverse of the number of local branches) whenever each of its branches is specialized in a different industry. We then look at the distribution of this index across banks. Results are shown on the last line of each panel in Table 2.3. We find that this HHI index is low on average (0.15), with a value of 0.1 for more than 75%s of larger, "network" banks. This confirms that, within-bank, branch specialization is barely concentrated in a few industries only.

4. Empirical strategy

4.1. Branch closures and credit supply: firm-bank level analysis

We first aim to identify the loan supply pattern of banks to existing corporate clients around the closing date of bank branches serving these customers. For this purpose, we estimate the change in credit experienced by firm i that borrows from a bank b when the firm is being transferred from a closing branch h to an absorbing branch h' within bank b. We aggregate granular firm-branch credit exposures at the firm-bank level and estimate the following empirical model :

$$\begin{aligned} \text{Log(Loan)}_{ibt} = \; & \beta_0 \mathbb{I}\{\text{Announcement}\}_{ib,T-6\leq t\leq T} + \beta_1 \mathbb{I}\{\text{Closing}\}_{ib,t\geq T} \\ & + \text{Firm x Bank FE} + \text{Firm x Quarter FE} \\ & + \text{Banking group x Quarter FE} + \epsilon_{ijt} \end{aligned} \tag{2.4}$$

where i denotes a firm, b denotes a bank and $t = T$ is the quarter of closing. The dependent variable $Loan_{ibt}$ measures the euro amount of total debt outstanding between firm i and bank b's branch(es) in quarter t. The indicator variable $\mathbb{I}\{\text{Announcement}\}_{b,T-6\leq t\leq T}$ is set to one during the transition period that follows the announcement of the closing, six quarters before the branch is definitively closed. The indicator variable $\mathbb{I}\{\text{closing}\}_{b,t\geq T}$ takes the value of one after date $t = T$ when the branch of b that used to lend to i definitively closes (meaning that firm i had a positive credit exposure with the closed branch at least one quarter during the 12 quarters that precede the official closing).

We estimate (2.4) over the period from Q1 2010 to Q3 2017. Our main coefficient of interest is β_1 which captures to what extent banks that experienced a branch closing modify their loan supply when borrowers are reallocated between branches. As said above, we restrict the sample to firms located in urban units where at least one bank branch closed over the sample period. This restriction is intended to both limit discrepancies between geographical areas that are treated by branch closures and geographical areas that are not and to reduce the imbalance in terms of size between of our treatment and control groups.

In estimating (2.4), we face to identification issues. The first one is a possible endogeneity of the bank's choice to close a branch to the condition of local borrowers: a closing may take place in a given area because, e.g., of the lackluster profitability of loans to local firms,

reflecting a decline in economic activity in this neighborhood. We address this concern by including firm-bank fixed effects that control for time-invariant characteristics linked to the respective firm-bank match.

Second, we also face the standard challenge of disentangling demand and supply of credit. For this purpose, we include firm-time fixed effects in our regression, which absorb all possible shocks to firms' demand and credit quality, in the spirit of Khwaja and Mian (2008) and many others. This implies that we consider only firms that borrow from at least two banks. We therefore identify the effect of bank branch closures "within" the borrowing firm, i.e., we compare credit amounts borrowed by the same firm in a given quarter from two different banks (typically one that closes its local branch and one that does not), before and after the branch closing that affects one of the credit relationships maintained by the firm. Last, note that we also include banking group-quarter fixed effects among the regressors in (2.4) in order to control for shocks at the more aggregate banking-group level that may also govern local lending decisions.

4.2. Branch closures and credit supply: firm-level analysis

In a second step, we check whether a firm is able to compensate for the negative credit shock that may be associated with a branch reallocation within one of its lenders. For this purpose, we collapse our database at the firm level and estimate the following model:

$$\begin{aligned}\text{Log(Loan)}_{it} = {} & \beta_0 \mathbb{I}\{\text{Announcement}\}_{i,T-6\leq t\leq T} + \beta_1 \mathbb{I}\{\text{closing}\}_{i,t\geq T} \\ & + \text{Firm FE} + \text{Industry x quarter FE} \\ & + \text{Banking group x Urban unit x quarter FE} + \epsilon_{it}\end{aligned} \tag{2.5}$$

where i stands for the firm. Now, the dummy variable $\mathbb{I}\{Closing\}_{t\geq k}$ takes the value of one whenever the firm faces the closing of one of its relationship bank branches. Again, we control for unobserved characteristics of the firms, as well as for demand shocks and time varying local factors at the industry level and geographical location level, as well as for other unspecified bank-level shocks, by saturating the model with appropriate sets of fixed effects.

4.3. Branch closures and change in the industry specialization of the bank

Last, we investigate whether and to what extent local bank specialization contributes to explaining the effect of a branch reallocation on the supply of credit to the bank's customers.

We test for the role of the local specialization of bank branches by estimating an augmented version of the previous empirical model (2.4):

$$
\begin{aligned}
\text{Log(Loan)}_{ibt} = & \ \beta_0 \mathbb{I}\{\text{Announcement}\}_{ib,T-6\leq t\leq T} + \beta_1 \mathbb{I}\{\text{Closing}\}_{ib,t\geq T} \\
& + \beta_2 \Delta S_{b,i,u} \mathbb{I}\{\text{Announcement}\}_{ib,T-6\leq t\leq T} \\
& + \beta_3 \Delta S_{b,i,u} \mathbb{I}\{\text{Closing}\}_{ib,t\geq T} \\
& + \beta_4 S_{b,i,u} \\
& + \text{Firm x Bank FE} + \text{Firm x Quarter FE} \\
& + \text{Banking group x Quarter FE} + \epsilon_{ijt}
\end{aligned}
\tag{2.6}
$$

where $\Delta S_{b,i,u}$ measures the *relative loss* in *firm-specific* branch specialization when the firm i is reallocated from the closed branch h to another branch h' of the same bank b. For instance, $\Delta S_{b,i,u}$ takes the value of +1 whenever the closed branch was specialized in lending to the firm's industry $s(i)$, whereas the new branch is not. We then control for the initial level of borrower-specific industrial specialization of the closed branch and saturate the model with fixed effects in order to control for firm-bank matching and for changes in credit demand.

As above, we are interested in evaluating whether the average small firm is able to compensate for the negative impact of the bank specialization loss it is subjected to by borrowing more from other lenders. We therefore run a similar regression after having collapsed the data at the firm level:

$$
\begin{aligned}
\text{Log(Loan)}_{it} = & \ \beta_0 \mathbb{I}\{\text{Announcement}\}_{i,T-6\leq t\leq T} + \beta_1 \mathbb{I}\{\text{Closing}\}_{i,t\geq T} \\
& + \beta_2 \Delta S_{i,u} \mathbb{I}\{\text{Announcement}\}_{i,T-6\leq t\leq T} \\
& + \beta_3 \Delta S_{i,u} \mathbb{I}\{\text{Closing}\}_{i,t\geq T} \\
& + \beta_4 S_{i,u} \\
& + \text{Firm FE} + \text{Industry x quarter FE} \\
& + \text{Banking group x Urban unit x quarter FE} + \epsilon_{it}
\end{aligned}
\tag{2.7}
$$

where $\Delta S_{i,u}$ denotes the bank specialization loss faced by firm i because of a branch reallocation within one of its banks at time T.

5. Results

5.1. *Branch reallocation and credit supply at the firm-bank level*

Table 2.4 shows our results when we estimate the effect of a branch closing on credit at the firm-bank level, as specified in equation (2.4). We find that, within their bank, firms that are reallocated to a new branch after a branch closing face a sizeable reduction in credit supply. A proper identification of this negative supply shock may be hampered if the decision to close the branch is correlated with other factors potentially also affecting the new branch, such as local economic conditions, the average credit quality of borrowers, regulatory constraints etc. To address such concerns, in columns 1 to 4 of Table 2.4, we progressively saturate our model with an extensive set of fixed effects.

First, in column (1), thanks to the inclusion of bank-firm fixed effect among regressors, we account for potential differences in borrower-lender pairs, e.g. weak banks matching with weak borrowers. This implies that our effect is estimated within an existing bank-firm relationship (intensive margin), and does not capture the creation or destruction of bank-firm relationships (extensive margin). Next, in column (2), we control for firm-level demand (or credit quality) shocks by including firm-quarter fixed effects. We then add banking group-quarter fixed effects to account for fluctuations in credit supply that would be driven by shocks at the group level (such as regulatory shocks on the level of required capital. Last, we control for variations of banking groups' credit supply at the local level (such as a strategic decision to divest from some areas) by including banking group-urban unit-quarter dummies.

Whatever the set of fixed effects used as controls, we find robust evidence of a significant drop in the amount of credit supplied to a reallocated firm. The credit contraction is already significant during the transition period after the branch closing announcement but before the effective closing and transfer of the firm's account. The magnitude of the contraction in credit supplied is however twice as large after the reallocation. This contraction is economically significant: the amount of total credit made available to the firm (including undrawn credit lines) drops by some 12% due to the branch reallocation. For the average firm-bank relationship in our sample, this amounts to a cut by some $60,000$ euros.

In the last two columns of the table, we provide complementary results on the impact of a branch reallocation on the credit mix supplied to the firm and the distance between borrower and lender. Column (5) shows that the ratio of long-term credit over total credit (drawn and undrawn) is not significantly affected by the reallocation. This suggests that firms facing a

branch closing do not draw more on credit lines thereafter. Column (6) confirms the intuition that a branch reallocation increases the average distance between a borrower and its bank. The estimated coefficient corresponds to a doubling firm-bank distance for reallocated firms.

Last, we take a closer look at how the identified effect unfolds over time. Figure 2.7 shows the estimation results of a dynamic version of equation 2.4, where we interact the closing indicator with dummies for each quarter instead of a unique ex-post period.[19] Credit supply starts to drop around closing announcement. After the closing, the contraction in lending is quite persistent and holds out over at least three years. The result is consistent with Nguyen (2019) who shows that, following a merger, branch closures lead to a decline in local small business lending that persists during up to six years. Importantly for identification, the figure also shows that the usual parallel trend assumption is vindicated before the closing announcement.

5.2. *Can multi-bank firms compensate for the effect of a branch reallocation?*

Table 2.5 shows estimation results for equation (2.5), when the bank-firm data is collapsed at the firm level. Note that we use here the same sample of firms as in the analysis above. We find that the average small firm borrowing from at least two banks is only partly able to compensate for the decline in credit that follows a branch reallocation within one of its banks. More precisely, the branch reallocation seems to have no effect on the total supply of credit to the firm during the transition period, while the ratio of long-term drawn lending to total credit slightly decreases. This suggests that the firm may be able to negotiate in the short run an increase of its credit lines with the other lenders. However, once the branch closing is effective, the average small firm still faces a contraction by some 4% of the total of its bank borrowings.

Last, figure 2.8 shows the estimated coefficients of interest in a dynamic version of (2.5). The total amount of credit borrowed by a treated firm reaches a trough after two years. The contraction in the firm's access to credit is still significant three years after a branch reallocation.

5.3. *The benefits of the local industry specialization of banks*

We now turn to estimating equations (2.6) and (2.7), so as to highlight how the local industry specialization of banks shapes bank lending to small firms. First, table 2.6 presents the results

[19] In the figure, we take as a reference period the quarter two years-ahead of the effective closing (i.e., the corresponding interacted dummy is excluded from the regressors). Results are unchanged if we take the whole pre-announcement period as a reference.

of regression (2.6) using firm-bank data. Columns (2) to (4) show robust evidence that the drop in credit supply associated with a branch reallocation is strongly amplified when the firm looses the benefit of being connected to a bank branch that was specialized in its industry. Small firms that are matched to a specialized branch receive on average between 6% and 10% more credit on a regular basis. When a branch reallocation entails a bank specialization loss, the contraction in credit supply is roughly twice as large as otherwise (with a coefficient for the interacted term pointing to a cut in outstanding amounts by between -9% and -13%). As shown in column (5), the effect is not symmetric: borrowers that end up being matched to a branch with relevant specialization remain negatively affected by the branch reallocation in spite of the gain in terms of skills of the lender. Table 2.7 presents additional results at the firm level (cf. equation 2.7). We find again that a firm which looses access to a branch with the relevant industry specialization within one bank cannot fully compensate for the lower provision of credit by borrowing more from its other lenders. The aggregate effect is however less significant than the effect we observe within the reallocating bank alone.

The impact of branch closures on credit supply to SMEs may also go through other channels which may be correlated with changes in industry specialization. We therefore check for the robustness of our findings to the inclusion of additional interaction terms in regression (2.6). Table 2.8 shows our results. First, we find that the effect of the loss in bank specialization is not wiped out by the concomitant increase in bank-firm distance due to the branch reallocation. Specifically, we interact in column (2) our closing dummy *closing* $\times$ *post* with an indicator variable $\Delta distance$ which equals 1 (resp. -1) if the closing entails an increase (resp. a decrease) in the firm-bank distance larger than 20 kilometers, i.e. twice the median of the firm-branch distance in our whole sample. We observe that the increased distance does not seem to matter much (the coefficient is negative as expected but non significant). Accordingly, the specialization coefficient does not bulge.

Second, we test whether the coefficient for the loss in bank specialization captures the possible confounding effect of a decrease in local lending concentration associated with the branch reallocation. In column (3), we interact the closing event with a dummy variable equals to one if local bank competition (measured in the urban unit of the absorbing branch) is lower than the median competition. The coefficient is negative and significant (-0.09), which suggests that entering an environment of weak bank competition worsens the effect of branch reallocation. However, controlling for changes in competition in the local credit market due to the transfer does not wipe out the effect of the specialization loss *per se*.

Last, we further investigate the role of local bank competition in shaping the response of

SME lending to branch closures. For this purpose, we divide our sample of closing events into two groups: closures with an absorbing branch that is located in a *low-competition* urban unit, in column (5) versus located in a *high-competition* urban unit, in column (6). In an environment of low competition, a branch closing hits homogeneously all transferred firms, whatever the degree of specialization of the closed branch. A rationale for this could be that less effort is required from the bank in the new environment for keeping the customer and reaping the rents associated with prior information extraction. On the contrary, branch specialization and the associated loss in industry-specific knowledge following to a closing matters in a highly competitive banking environment: the coefficient for the closing *per se* is not significant anymore, except for firms that additionally are subjected to a bank specialization loss. This suggests that the reduced supply of credit is then merely a consequence of the loss in loan officers' skills required to accommodate the credit needs of these customers. In the standard case however, a high competitive pressure is enough to ensure continuity in the level of credit supplied by the reallocating bank to small firms.

6. Conclusion

In this paper, we provide evidence that a large share of local bank branches in France are specialized in lending to some industries. We also show that this local bank specialization matters for SMEs' access to credit: a small firm enjoys on average a better access to bank credit when the local branch of its relationship bank has gained a better knowledge of the industry to which the firm belongs. Using a very rich dataset on bank branch closures over 2010-2017, we find that closures on average entail a substantial and persistent loss in credit supplied for the local customers which are transferred to another branch. Importantly however, firms which, as a consequence, lose access to loan officers specialized in their industry are more badly hit than others. This effect is not wiped out when considering confounding factors such as the increased distance or the decreased local bank competition that may be associated with the transfer.

Our findings have possibly important policy implications. First, competition regulators overseeing cuts in bank branch networks following bank mergers should be aware that local bank specialization exists and matters for credit supply to SMEs. Concerns may arise if simple competition rules, such as the obligation to close one of two branches in a given area, are implemented bluntly. Second, supervisors and bank managers should be aware that closing specialized branches may have differentiated effects on small borrowers, depending on their

industry.

Many issues remain however open, and left for further research. A first, looming question is of course to understand how and why local bank branches become specialized in funding one specific industry (more than other neighbour branches do). A possible cause of specialization may be diffusion effects within networks of customers (e.g., buyers and suppliers, or members of local professional clubs): in other words, the lending pattern of a branch is likely to be history-dependent. An idiosyncratic role of some loan officers, who e.g. join a branch with a better experience in dealing with a certain type of firms, also comes to mind. In this case, the pending issue for bank managers is to ensure that such human capital is not lost whenever the branch closes. Drexler and Schoar (2014) suggest that soft information can be transferred between bank employees when loan officer turnover is properly anticipated. The same could be expected regarding transfers of industry-specific knowledge.

References

Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. *Review of Financial Studies*, 23(7):2757–2788.

Benetton, M. (2020). Leverage regulation and market structure: A structural model of the uk mortgage market. *The Journal of finance*, Forthcoming.

Berger, A. and Udell, G. (1995). Relationship lending and lines of credit in small firm finance. *The Journal of Business*, 68(3):351–81.

Berger, A. and Udell, G. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *Economic Journal*, 112(477):F32–F53.

Bonfim, D., Nogueira, G., and Ongena, S. (2016). Sorry, we re closed: Loan conditions when bank branches close and firms transfer to another bank. Working papers, Banco de Portugal, Economics and Research Department.

Boot, A. W. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, 9(1):7–25.

De Jonghe, O., Dewachter, H., Mulier, K., Ongena, S., and Schepens, G. (2019). Some borrowers are more equal than others: bank funding shocks and credit reallocation. Working Paper Series 2230, European Central Bank.

Degryse, H., Kim, M., and Ongena, S. (2009). *Microeconometrics of Banking Methods, Applications, and Results*. Oxford University Press.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *Journal of Finance*, 60(1):231–266.

Drexler, A. and Schoar, A. (2014). Do relationships matter? evidence from loan officer turnover. *Management Science*, 60(11):2722–2736.

Egan, M., Hortaçsu, A., and Matvos, G. (2017). Deposit competition and financial fragility: Evidence from the us banking sector. *American Economic Review*, 107(1):169–216.

Goedde-Menke, M. and Ingermann, P.-H. (2020). The impact of organizational downsizing on loan officer specialization and credit defaults. mimeo, WWU Muenster.

Hastings, J., Hortaçsu, A., and Syverson, C. (2017). Sales force and competition in financial product markets: The case of mexico's social security privatization. *Econometrica*, 85(6):1723–1761.

Ioannidou, V. and Ongena, S. (2010). "time for a change": Loan conditions and bank behavior when firms switch banks. *Journal of Finance*, 65(5):1847–1877.

Khwaja, A. I. and Mian, A. (2008). Tracing the impact of bank liquidity shocks: evidence from an emerging market. *American Economic Review*, 98:1413–1442.

Nguyen, H.-L. Q. (2019). Are Credit Markets Still Local? Evidence from Bank Branch Closings. *American Economic Journal: Applied Economics*, 11(1):1–32.

Ongena, S. and Smith, D. C. (2000). What determines the number of bank relationships? cross-country evidence. *Journal of Financial Intermediation*, 9(1):26–56.

Ongena, S. and Yu, Y. (2017). Firm industry affiliation and multiple bank relationships. *Journal of Financial Services Research*, 51(1):1–17.

Paravisini, D., Rappoport, V., and Schnabl, P. (2017). Specialization in Bank Lending: Evidence from Exporting Firms. (12156).

Petersen, M. A. and Rajan, R. G. (1994). The Benefits of Lending Relationships: Evidence from Small Business Data. *Journal of Finance*, 49(1):3–37.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of finance*, 47(4):1367–1400.

Sharpe, S. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance*, 45(4):1069–87.

Figures

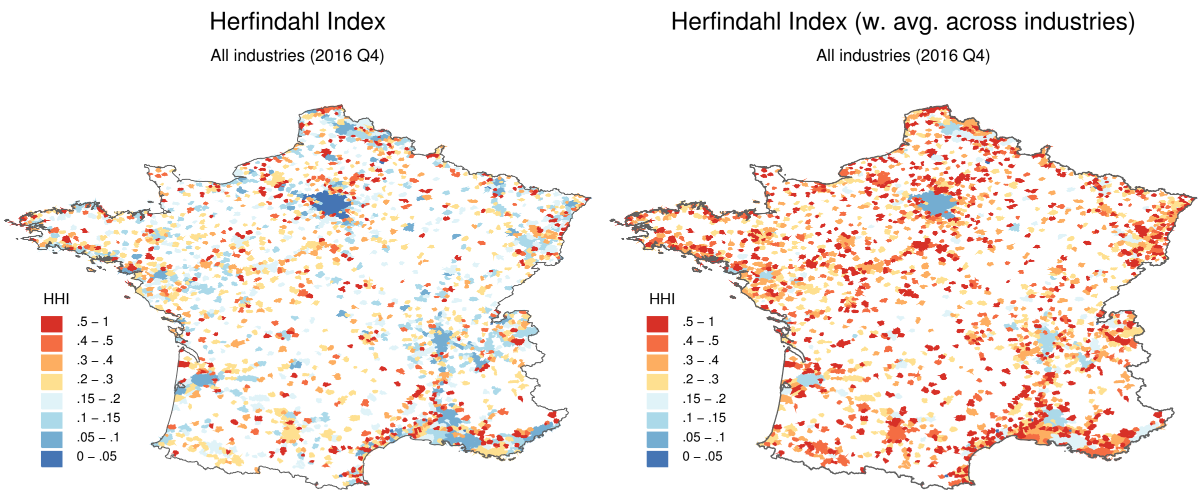


Fig. 2.1. Herfindahl Index at urban unit level

Note: The figure plots, for each urban unit, two different measures of credit market concentration. The first Herfindahl Index (left panel) is the standard concentration measure, calculated using total lending shares within each unit. The second (right panel) is the (credit) weighted average of industry-specific credit market concentrations within each unit, which therefore accounts for market segmentation. Data as of December 2016.

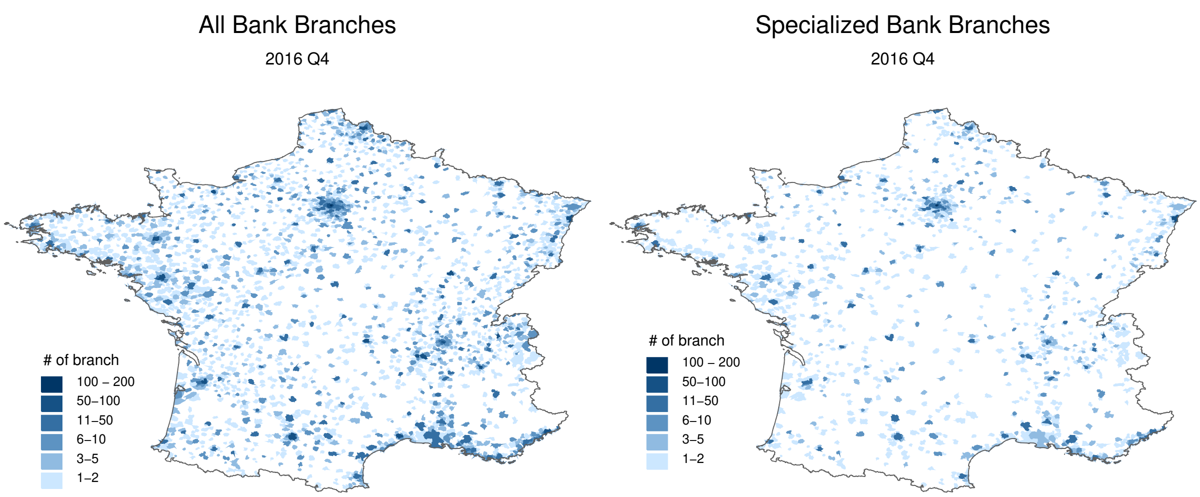


Fig. 2.2. Geographical distribution of bank branches lending to SMEs in France (2016).

Note: This map locates all bank branches lending to SMEs in mainland France as of December 2016 (left panel) vs "specialized" branches only (right panel). See section 3 for the definition of the industry specialization of bank branches. Location refers to the ZIP code of the branch. ZIP codes in white are cities without any bank branch lending to SMEs in December 2016. The darkest color stands for ZIP codes where more than 100 bank branches are located.

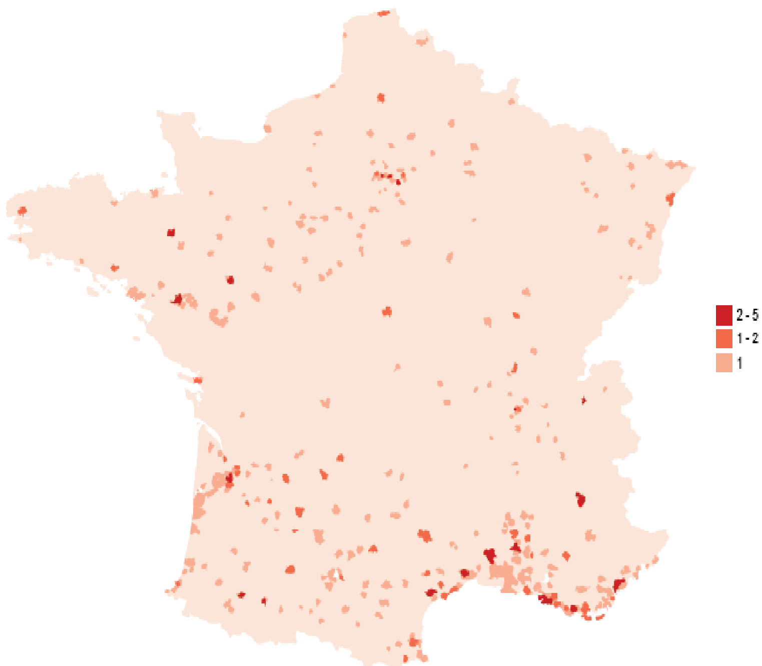


Fig. 2.3. The geography of bank branch closures in France, 2010-2017.

Note: This map locates bank branch closures in mainland France over the period from 2010 Q1 to 2017 Q3. Location refers to the ZIP code of the closed down branch. The sample is limited to branches actively lending to SMEs according to the credit register and, before 2015, to branches of banks that had already opted for the "RIB permanent"system of account management.

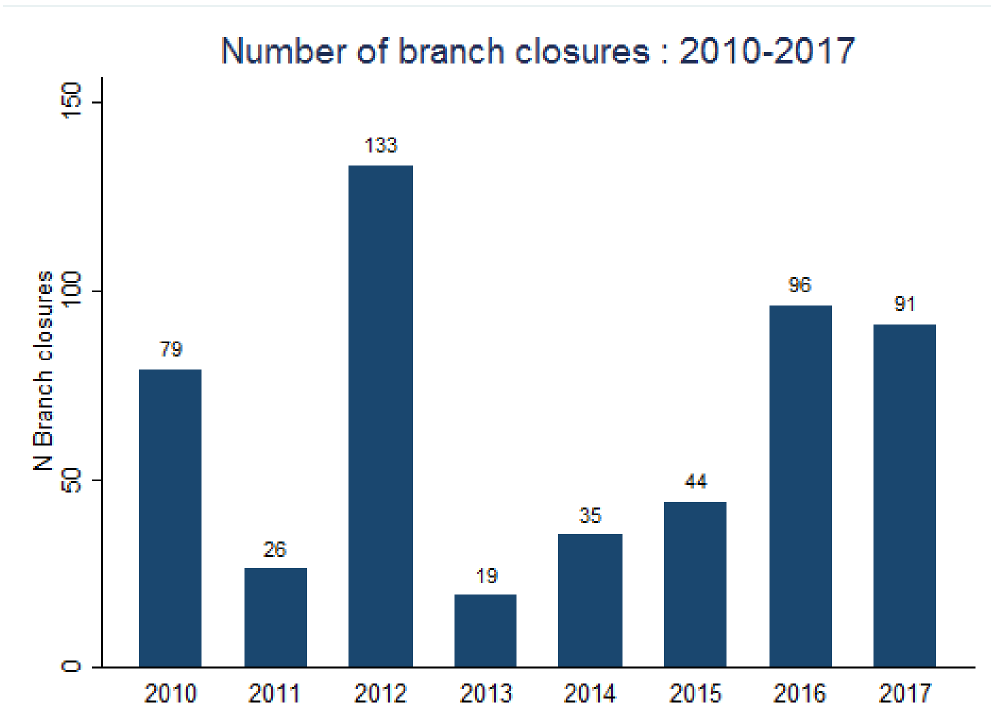


Fig. 2.4. Time series of branch closures in France, 2010-2017.

Note: This graph shows the annual number of branch closures over the period from 2010 Q1 to 2017 Q3 in our cleaned dataset. The sample is limited to branches actively lending to SMEs according to the credit register, and, before 2015, to branches of banks that had already opted for the "RIB permanent"system of account management.

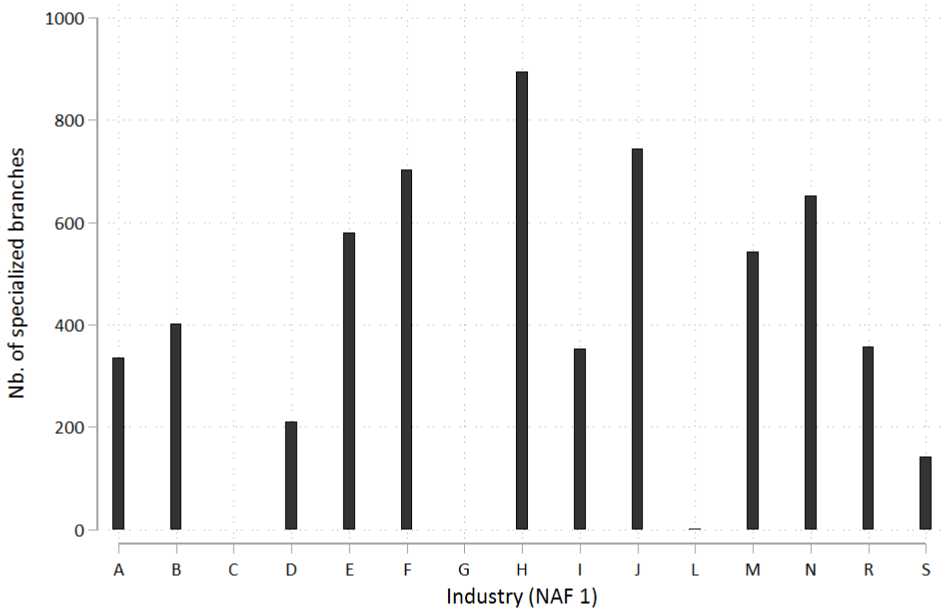


Fig. 2.5. Number of specialized bank branches by borrowing industry.

Note: This graph shows the number of bank branches specialized in each of the 15 NACE rev2 one-digit industries (among branches lending to SMEs in our sample). A: Farming, forestry and fishing, B: Mining and quarrying, C: Manufacturing, D: Electricity, gas, steam and air conditioning supply, E: Water supply, sewerage and waste management, F: Construction, G: Wholesale and retail trade and repair of motor vehicles, H: Transportation and storage, I: Accommodation and food services activities, J: Information and communication, L: Real estate activities, M: Professional, scientific and technical activities, N: Administrative and support service activities, R: Arts, entertainment and recreation, S: Other service activities.

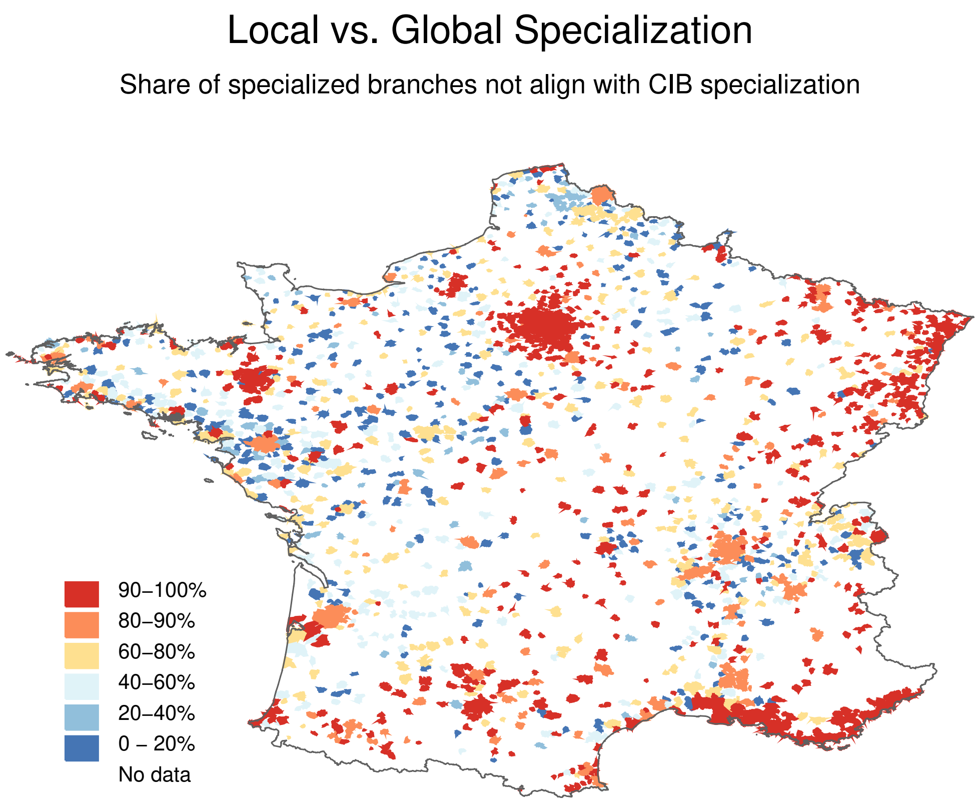


Fig. 2.6. Share of specialized branches not aligned with the industry specialization of their parent bank, by urban unit (2016).

Note: The figure plots, for each urban unit, the share of specialized bank branches with an industry specialization which differs from the industry specialization of their parent bank. Data as of December 2016.

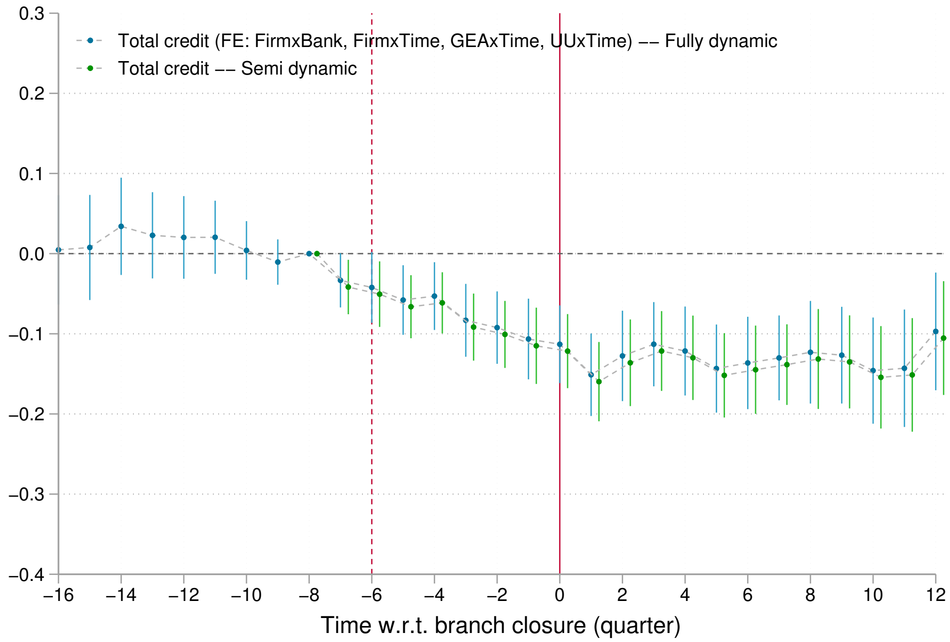


Fig. 2.7. Impact of a branch closure on credit: firm-bank level analysis.

Note: This graph shows estimation results for a dynamic version of equation 2.4. The dependent variable is the natural logarithm of total credit outstanding (in euro thousands) at the bank-firm level in quarter t, including unused credit lines. Time 0 is the date of the closure of the bank's branch which lends to i. Coefficients for each quarter, prior to, and following the bank branch closure, are plotted along with 95% confidence intervals. Quarter $t-8$ is used as a reference and the corresponding dummy is omitted from the regression. The time period between the dashed line and the solid line is the 6-quarter transition 's branch period between the announcement date of the branch closure and the official closing date. The sample period is Q1 2010 to Q3 2017.

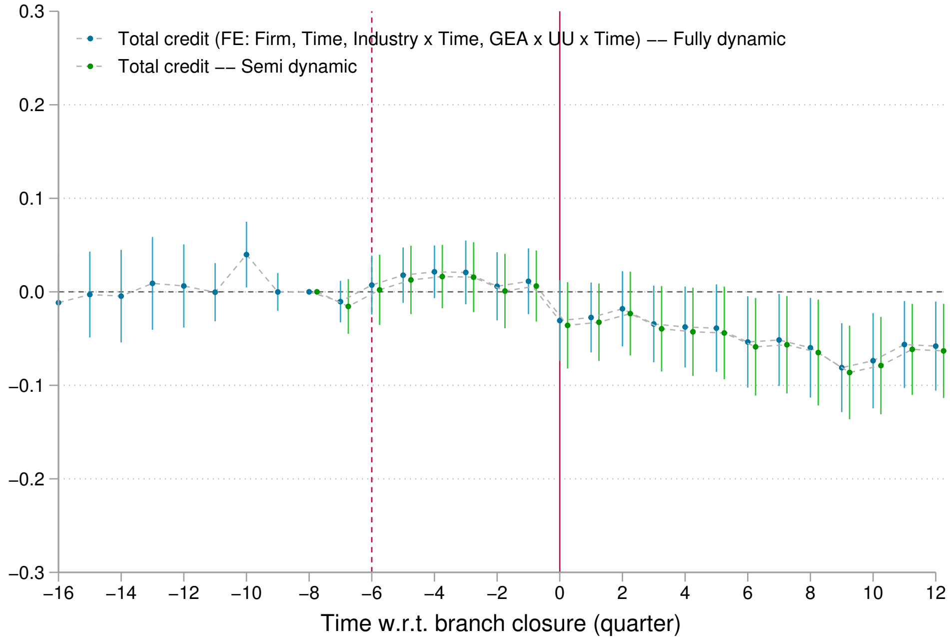


Fig. 2.8. Impact of a branch closure on credit: firm-level analysis.

Note: This graph shows estimation results for a dynamic version of equation 2.5. The dependent variable is the natural logarithm of total bank credit outstanding (in euro thousands) at the firm level in quarter t, including unused credit lines. Time 0 is the date of the closure of a bank branch which lends to i. Coefficients for each quarter, prior to, and following the bank branch closure, are plotted along with 95% confidence intervals. Quarter $t-8$ is used as a reference and the corresponding dummy is omitted from the regression. The time period between the dashed line and the solid line is the 6-quarter transition 's branch period between the announcement date of the branch closure and the official closing date. The sample period is Q1 2010 to Q3 2017.

Tables

Table 2.1: Firm-bank-level summary statistics.

	p10	p25	Median	p75	p90	Mean	Nb.Obs.
Firm-Bank x quarter level of obs.							
Nb. firms per quarter	72603.0	75997.0	78742.0	80111.0	80306.0	77520.4	5,211,791
Nb. banks per quarter	311.0	322.0	346.0	377.0	379.0	346.1	5,211,791
Nb. branches per quarter	10665.0	11246.0	12054.0	12317.0	12599.0	11780.3	5,211,791
Nb. firm-bank rel. per quarter	174200.0	185194.0	190519.0	192558.0	195465.0	187246.7	5,211,791
Total Loan	34.0	63.0	150.0	380.0	939.0	527.6	5,211,791
- Drawn credit	8.0	43.0	115.0	308.0	779.0	427.7	5,211,791
- Long-term loans	0.0	0.0	39.0	164.0	491.0	294.5	5,211,791
- Short-term loans	0.0	0.0	1.0	86.0	294.0	133.2	5,211,791
- Undrawn credit line	0.0	0.0	0.0	29.0	151.0	99.8	5,211,791
MLT ratio	0.0	0.0	0.5	1.0	1.0	0.5	5,211,791
Average distance in km							
Total Loan	0.0	4.3	13.6	56.2	272.0	75.3	5,211,791
- Distance, Drawn credit	0.0	4.2	12.8	51.4	209.9	69.0	4,811,504
- Distance, Long-term loans	0.0	3.6	10.9	41.2	115.4	51.3	3,354,248
- Distance, Short-term loans	0.0	4.7	14.4	58.6	289.0	77.9	2,648,396
- Distance, Undrawn credit line	0.0	5.1	17.1	72.9	351.4	88.2	1,701,646

Note: The sample period is Q1 2010 to Q3 2017. The average distance between a firm and its lenders, for each type of credit, is computed conditionally on credit being non null. All credit variables are in euro thousands. MLT ratio is the ratio of long-term credit over the total of drawn credit and unused credit lines.

Table 2.2: Firm-level summary statistics

	p10	p25	Median	p75	p90	Mean	Nb.Obs.
Firm x quarter level of obs.							
Total loan	63.0	141.0	339.0	865.0	2221.0	1243.8	1,408,070
- Drawn credit	32.0	95.0	261.0	700.0	1824.0	1008.1	1,408,070
- Long-term loans	0.0	21.0	114.0	378.0	1089.0	685.7	1,408,070
- Short-term loans	0.0	0.0	27.0	213.0	714.0	322.4	1,408,070
- Undrawn credit line	0.0	0.0	10.0	114.0	393.0	235.7	1,408,070
MLT ratio	0.0	0.1	0.5	0.9	1.0	0.5	1,408,070
Average distance in km							
Total loan	0.8	6.2	21.7	72.6	196.5	65.7	1,408,070
- Distance, Drawn Credit	0.0	6.0	20.3	66.9	187.3	63.3	1,408,070
- Distance, Long-term loans	0.0	4.8	14.6	45.1	113.9	47.1	1,408,070
- Distance, Short-term loans	0.0	6.1	21.2	84.8	230.7	76.9	1,408,070
- Distance, Undrawn credit line	0.0	6.1	23.4	98.4	302.6	89.6	1,408,070

Note: The sample period is Q1 2010 to Q3 2017. The average distance between a firm and its lenders, for each type of credit, is computed conditionally on credit being non null. All credit variables are in euro thousands. MLT ratio is the ratio of long-term credit over the total of drawn credit and unused credit lines.

Table 2.3: Summary statistics on the specialization of bank branches.

	All banks						
	Mean	Min	p25	Median	p75	Max	Nb.Obs.
Number of branches	33.95	1.0	1.0	1.0	20.0	1152.0	294
Number of specialized branches	11.65	0.0	1.0	1.0	8.0	390.0	294
Share of specialized branches	0.58	0.0	0.2	0.6	1.0	1.0	294
Industrial concentration of specialized branches	0.29	0.0	0.1	0.2	0.3	1.0	294
	Larger banks						
	Mean	Min	p25	Median	p75	Max	Nb.Obs.
Number of branches	119.40	11.0	39.0	68.0	107.0	1152.0	81
Number of specialized branches	39.83	4.0	10.0	21.0	41.0	390.0	81
Share of specialized branches	0.37	0.1	0.2	0.3	0.5	1.0	81
Industrial concentration of specialized branches	0.15	0.1	0.1	0.1	0.2	0.5	81

Note: For each quarter from Q1 2010 to Q3 2017, we compute the share of specialized branches as the ratio of the number of bank branches that are specialized in at least one industry over the total number of bank branches active in our final sample. The industrial concentration of specialized branches is an HHI index that is equal to one if, within a bank, all specialized branches are specialized in the same industry and close to zero if each specialized branch is specialized in a different industry. Larger banks are banks which operate more than 10 branches lending to SMEs.

Table 2.4: Branch closures and SMEs' access to credit: firm-bank level analysis.

	Total credit				MLT ratio	Distance
	(1)	(2)	(3)	(4)	(5)	(6)
Closing x Post	-0.252***	-0.112***	-0.124***	-0.116***	0.009	0.678***
	(0.012)	(0.018)	(0.019)	(0.021)	(0.006)	(0.043)
Announcement x Post	-0.143***	-0.059***	-0.072***	-0.064***	0.002	0.193***
	(0.010)	(0.015)	(0.016)	(0.017)	(0.005)	(0.037)
Firm x Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm x Quarter FE		Yes	Yes	Yes	Yes	Yes
Banking Group x Quarter FE			Yes		Yes	Yes
Banking Group x UU x Quarter FE				Yes		
Observations	5,177,193	3,932,607	3,931,733	3,821,431	3,931,733	3,821,431
R-Square	0.80	0.90	0.90	0.91	0.89	0.98

Note: This table shows estimation level for regression (2.4). The sample period is Q1 2010 to Q3 2017. The dependent variable in columns 1 to 4 is the log amount of total credit outstanding (in euro thousands) at the firm-bank i, b level, in quarter t (including unused credit lines). The dependent variable in column 5 is the ratio of medium to long-term loans over total loans for firm-bank relationship $i - j$ in quarter t. The dependent variable in column 6 is the natural logarithm of the distance in kilometers between a firm and its banks. *Closing* $\times$ *Post* is the 16-quarter period starting after the official date of branch closure. *Announcement* $\times$ *Post* is the 6-quarter period starting with the announcement of the branch closure and ending with the official branch closure. *Urban Unit* denotes the urban unit of the location of the firm. Robust standard errors (clustered at the urban unit x year level) are in parentheses.

Table 2.5: Branch closures and SMES' access to credit: firm-level analysis

	Total credit					MLT ratio	Distance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Closing × Post	-0.050***	-0.041**	-0.041**	-0.039**	-0.041**	-0.005	0.115***
	(0.018)	(0.019)	(0.019)	(0.019)	(0.020)	(0.005)	(0.026)
Announcement × Post	-0.003	0.011	0.014	0.016	0.016	-0.007*	0.057
	(0.012)	(0.013)	(0.013)	(0.013)	(0.014)	(0.004)	(0.022)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes						
Urban Unit x Quarter FE		Yes	Yes	Yes		Yes	Yes
Industry x Quarter FE			Yes	Yes	Yes	Yes	Yes
Banking Group x Quarter FE				Yes		Yes	Yes
Banking Group x UU x Quarter FE					Yes		
N of clusters	726	726	726	726	726	726	726
Observations	1,408,070	1,408,070	1,408,070	1,408,070	1,408,070	1,408,070	1,408,070
R-Square	0.84	0.84	0.85	0.85	0.86	0.78	0.84

Note: This table shows estimation level for regression (2.6). The sample period is Q1 2010 to Q3 2017. The dependent variable in columns 1 to 4 is the log amount of total credit outstanding (in euro thousands) at the firm i level, in quarter t (including unused credit lines). The dependent variable in column 5 is the ratio of medium to long-term loans over total loans for firm i in quarter t. The dependent variable in column 6 is the natural logarithm of the distance in kilometers between a firm and its banks. *Closing* $\times$ *Post* is the 16-quarter period starting after the official date of branch closure. *Announcement* $\times$ *Post* is the 6-quarter period starting with the announcement of the branch closure and ending with the official branch closure. *Urban Unit* denotes the urban unit of the location of the firm. Robust standard errors (clustered at the urban unit x year level) are in parentheses.

Table 2.6: Branch closures, branch specialization and SMEs' access to credit: firm-bank level analysis.

	Total credit						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Announcement x Post	-0.072***	-0.071***	-0.071***	-0.069***	-0.071***	-0.070***	-0.070***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Announcement x Post x Δ specialization		-0.053		-0.057			
		(0.050)		(0.050)			
Closing x Post	-0.124***	-0.118***	-0.121***	-0.116***	-0.112***	-0.115***	-0.111***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.019)
Closing x Post x Δ specialization		-0.174***		-0.145**		-0.090**	
		(0.058)		(0.057)		(0.046)	
Closing x Post x Specialization loss					-0.143**		-0.120**
					(0.057)		(0.057)
Closing x Post x Specialization gain					0.068		0.052
					(0.070)		(0.070)
Branch specialization (MA over 4Q)			0.062***	0.061***	0.061***		
			(0.006)	(0.006)	(0.006)		
Branch specialization (MA over 8Q)						0.109***	0.108***
						(0.008)	(0.008)
Firm x Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Banking Group x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,931,733	3,931,733	3,931,733	3,931,733	3,931,733	3,931,733	3931733
R-Square	0.90	0.90	0.90	0.90	0.00	0.90	0.90

Note: This table shows estimation level for regression (2.6). The sample period is Q1 2010 to Q3 2017. The dependent variable is the log amount of total credit outstanding (in euro thousands) at the firm-bank i, b level, in quarter t (including unused credit lines). *Closing* × *Post* is the 16-quarter period starting after the official date of branch closure. *Announcement* × *Post* is the 6-quarter period starting with the announcement of the branch closure and ending with the official branch closure. Δ *specialization* is the difference between the lagged level of specialization of the closing branch over 4 quarters (moving average) and the lagged level of specialization of the absorbing branch the quarter before the absorption. *Closing* × *Post* × *Specialization loss* indicates that the closing branch was specialized while the absorbing branch is not (Δ specialization = 1). On the contrary, *Closing* × *Post* × *Specialization gain* indicates that the closing branch was not specialized while the absorbing branch is (Δ specialization = -1). *Urban Unit* denotes the urban unit of the location of the firm. Robust standard errors (clustered at the urban unit x year level) are in parentheses.

Table 2.7: Branch closures, branch specialization and SMEs' access to credit: firm-level analysis.

	Total credit					
	(1)	(2)	(3)	(4)	(5)	(6)
Announcement × Post	0.018	0.018	0.018	0.018	0.019	0.036**
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.015)
Announcement × Post × Δ Specialization				0.044**		
				(0.017)		
Closing × Post	-0.039**	-0.039**	-0.038**	-0.039**	-0.023	-0.013
	(0.019)	(0.019)	(0.019)	(0.019)	(0.021)	(0.022)
Closing × Post × Δ Specialization			-0.044*			
			(0.026)			
Closing × Post × Specialization gain					0.062**	0.075*
					(0.028)	(0.040)
Closing × Post × Specialization loss					-0.081***	-0.116***
					(0.028)	(0.038)
Average bank specialization		-0.005	-0.005	-0.005	-0.006	-0.006
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Banking Group x UU x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,428,048	1,428,048	1,428,048	1,428,048	1,428,048	1,428,048
R-Square	0.85	0.85	0.85	0.85	0.85	0.85

Note: This table shows estimation level for regression (2.7). The sample period is Q1 2010 to Q3 2017. The dependent variable is the log amount of total credit outstanding (in euro thousands) at the firm i level, in quarter t (including unused credit lines). *Closing* × *Post* is the 16-quarter period starting after the official date of branch closure. *Announcement* × *Post* is the 6-quarter period starting with the announcement of the branch closure and ending with the official branch closure. Δ *specialization* is the difference between the lagged level of specialization of the closing branch over 4 quarters (moving average) and the lagged level of specialization of the absorbing branch the quarter before the absorption. *Closing* × *Post* × *Specialization loss* indicates that the closing branch was specialized while the absorbing branch is not (Δ specialization = 1). On the contrary, *Closing* × *Post* × *Specialization gain* indicates that the closing branch was not specialized while the absorbing branch is (Δ specialization = -1). *Urban Unit* denotes the urban unit of the location of the firm. Robust standard errors (clustered at the urban unit x year level) are in parentheses.

Table 2.8: Branch closures, branch specialization and SMEs' access to credit: the role of distance and competition.

	Total credit					
					Low comp.	High comp.
	(1)	(2)	(3)	(4)	(5)	(6)
Announcement x Post	-0.071***	-0.071***	-0.071***	-0.071***	-0.066***	-0.069
	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)	(0.048)
Closing x Post	-0.112***	-0.100***	-0.072***	-0.060*	-0.078**	-0.108
	(0.019)	(0.032)	(0.023)	(0.036)	(0.037)	(0.076)
Closing x Post x Specialization loss	-0.143**	-0.143**	-0.149***	-0.149***	-0.198***	-0.140
	(0.057)	(0.057)	(0.057)	(0.057)	(0.072)	(0.108)
Closing x Post x Specialization gain	0.068	0.068	0.080	0.080	-0.151*	0.344**
	(0.070)	(0.070)	(0.069)	(0.069)	(0.087)	(0.150)
Closing x Post x Δ distance		-0.013		-0.013	-0.039	0.086
		(0.028)		(0.029)	(0.034)	(0.056)
Closing x Post x Low competition			-0.086***	-0.086***		
			(0.029)	(0.029)		
Branch specialization (MA over 4Q)	0.061***	0.061***	0.061***	0.061***	0.051***	0.059***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.010)	(0.011)
Firm x Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Banking Group x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,931,733	3,931,733	3,931,733	3,931,733	1,920,910	1,408,592
R-Square	0.90	0.90	0.90	0.90	0.90	0.91

Note: This table shows estimation level for regression (2.6). The sample period is Q1 2010 to Q3 2017. The dependent variable is the log amount of total credit outstanding (in euro thousands) at the firm-bank i, b level, in quarter t (including unused credit lines). *Closing* × *Post* is the 16-quarter period starting after the official date of branch closure. *Announcement* × *Post* is the 6-quarter period starting with the announcement of the branch closure and ending with the official branch closure. Δ *specialization* is the difference between the lagged level of specialization of the closing branch over 4 quarters (moving average) and the lagged level of specialization of the absorbing branch the quarter before the absorption. *Closing* × *Post* × *Specialization loss* indicates that the closing branch was specialized while the absorbing branch is not (Δ specialization = 1). On the contrary, *Closing* × *Post* × *Specialization gain* indicates that the closing branch was not specialized while the absorbing branch is (Δ specialization = -1). *Urban Unit* denotes the urban unit of the location of the firm. Δ *distance* is a dummy variable equal to 1 if the closure triggers an increase in the firm-bank distance larger than 20 km, and -1 if the closure reduces the distance by more than 20 km. *Low competition* is a dummy equal to one if competition in the local credit market (measured in the UU of the absorbing branch) is below the median. *High competition* is a dummy equal to one if local bank competition is above the median. Robust standard errors (clustered at the urban unit x year level) are in parentheses.

A. Appendices

Table 2.9: Herfindhal Index at urban unit level

	Mean	p10	p25	Median	p75	p90
HHI (Traditionnal manner)	0.30	0.12	0.16	0.23	0.36	0.57
HHI (w. avg. across industries)	0.45	0.21	0.28	0.39	0.56	0.82

Note: This table shows two different measures of credit market concentration by urban unit, in 2016 Q4. The first Herfindahl Index (first line) is standard concentration measure, calculated using total lending shares. The second is the weighted average of credit concentrations calculated industry-by-industry (second line), which takes into account market segmentation.

Table 2.10: What drives bank branch closures ? (1) ANOVA.

	County FE	Urban Unit FE	Banking Group FE	Bank FE	(4) + County FE	(4) + UU FE
	(1)	(2)	(3)	(4)	(5)	(6)
R-Square	0.03	0.08	0.02	0.32	0.32	0.35
Explained by Bank FE					0.296	0.276
Explained by County FE					0.001	
Explained by UU FE						0.042

Note: The sample is the population of bank branches active in our cleaned dataset. The dependent variable is a dummy equals to 1 if the branch is closed during the 2010-2017 period, and 0 if not. We run a partial analysis of variance (ANOVA; Fixed-effect model class I) sequentially adding bank, banking group, department and urban unit fixed-effects. We report the R^2 and the share of variance explained by each category

Table 2.11: What drives bank branch closures ? (2) Linear probability model (department level).

	$\mathbb{P}$ (Branch closure over the period 2010-2017)					
	(1)	(2)	(3)	(4)	(5)	(6)
Bank branch size (Total Loan)	0.002***			0.002	0.002	0.003
	(0.001)			(0.002)	(0.002)	(0.002)
Market share (Branch total loan / County total loan)	-0.033			-0.033	-0.006	-0.022
	(0.050)			(0.052)	(0.052)	(0.047)
Dummy (Only branch of the bank in the county)	-0.018**			-0.018*	-0.021**	0.006
	(0.009)			(0.010)	(0.010)	(0.011)
Branch rank in the bank x county		-0.006		0.002	0.009	-0.016
		(0.010)		(0.011)	(0.011)	(0.010)
Branch rank in the bank		-0.005		-0.002	-0.009	0.024*
		(0.010)		(0.013)	(0.013)	(0.014)
County size (Population)			0.006***		0.007***	0.001
			(0.002)		(0.002)	(0.002)
County size variation: 1990-2006			0.053***		0.057***	-0.009
			(0.016)		(0.016)	(0.016)
Banking Group FE	✓	✓	✓	✓	✓	×
Bank FE	×	×	×	×	×	✓
Observations	15863	16084	16084	15863	15863	15722
R-Square	0.02	0.02	0.02	0.02	0.02	0.31

Note: The sample is the population of bank branches active in our cleaned dataset. The dependent variable is a dummy equals to 1 if the branch is closed during the 2010-2017 period, and 0 if not. We run a linear probability model with all our explanatory variables measured before the period of interest (i.e. in 2008 or 2009 for branch-level variables and 2006 for department level variables -taken from the French Census-). The *rank* of a branch is the rank of the branch (with respect to its size) within the department or within the department and the bank. It is equal to 1 if the branch is the largest in the department (respectively department x bank) and to N if the branch is the smallest of N branches located in this department (respectively department x bank). Total credit and population are in log.

Table 2.12: What drives bank branch closures ? Linear probability model (Urban unit level).

	$\mathbb{P}$ (Branch closure over the period 2010-2017)					
	(1)	(2)	(3)	(4)	(5)	(6)
Bank branch size (Total Loan)	0.001**			0.002	0.002	-0.000
	(0.001)			(0.001)	(0.001)	(0.002)
Market share (Branch total loan / UU total loan)	0.001			-0.001	-0.001	0.019***
	(0.007)			(0.007)	(0.007)	(0.006)
Dummy (Only branch of the bank in the UU)	-0.018***			-0.023***	-0.023***	-0.024***
	(0.003)			(0.005)	(0.005)	(0.004)
Branch rank in the bank x UU		-0.019***		0.011	0.012	-0.001
		(0.005)		(0.007)	(0.008)	(0.006)
Branch rank in the bank		0.004		-0.005	-0.005	-0.007
		(0.005)		(0.009)	(0.009)	(0.012)
UU size (Population)			0.003***		0.000	-0.003***
			(0.001)		(0.001)	(0.001)
UU size variation: 1990-2006			0.005		0.006	-0.004
			(0.005)		(0.005)	(0.005)
Banking Group FE	✓	✓	✓	✓	✓	×
Bank FE	×	×	×	×	×	✓
Observations	15,863	16,084	16,084	15,863	15,863	15,722
R-Square	0.02	0.02	0.02	0.02	0.02	0.32

Note: The level of observation is a bank branch. The sample is the entire population of bank branches active in our final sample. The dependent variable is a dummy equals to 1 if the branch undergoes a closure during the 2010-2017 period, and 0 if not. We run a linear probability model with all our explanatory variables measured before the period of interest (i.e. in 2008 or 2009 for branch-level variables and 2006 for urban-unit level variables). The branch rank measure the position of the branch (with respect to its size) within the urban unit or within the urban unit and the bank. The rank is equal to 1 if the branch is the largest one in the urban unit (respectively urban unit x bank) and to N if the branch is the smallest of N branches located in this urban unit (respectively urban unit x bank). Total credit and population are in log.

Table 2.13: Specialization, bank-firm distance and bank branch closures.

	Mean	Min	p5	p25	Median	p75	p95	Max
Δ specialization	0.1	-1.0	0.0	0.0	0.0	0.0	1.0	1.0
Δ distance (km)	23.1	-470.8	-22.7	2.4	16.0	42.2	76.2	379.7

Note: For each branch closing occurring between Q1 2010 and Q3 2017, we compute Δ spec. as the level of specialization of the closing branch minus the level of specialization of the absorbing branch, i.e. Δ spec. $= S_{b,s,u} - S_{b',s,u}$ where branch b is the closing one and branch b' the absorbing one (see 3.1). By definition, Δ spec. lies between -1 and 1. Similarly, we compute Δ distance as the distance from the closing branch minus the distance from the absorbing one, in km.

3

Aggregate Implications of Credit Relationship Flows

This chapter is based on a paper co-authored with Yasser Boualam (University of North Carolina - Chapel Hill).

Abstract

This paper studies the aggregate properties of credit relationship flows within the commercial loan market in France from 1998 through 2017. Using detailed bank-firm level data from the French Credit Register, we derive a novel decomposition for credit dynamics and show that banks actively and continuously adjust their lending supply along both extensive and *intensive margins. We document that gross flows associated with credit relationships (i) are volatile and pervasive throughout the cycle, and (ii) can account for up to 46% of the cyclical and 90% of the long-run variations in aggregate bank credit. We also highlight the role of the extensive margin in the transmission of monetary policy and show that the timing and magnitude of its response differ from that of the intensive margin.*

Keywords: Credit Flows; Financial Institutions; Monetary Policy Transmission; Relationship Lending; Search and Matching.

1. Introduction

What drives the fluctuations of credit over the business cycle and in the long run? How do banks adjust their credit supply in response to aggregate shocks or policy changes? These questions have been at the forefront of macro-finance and banking research at least since the seminal work of Bernanke (1983). Yet, our understanding of aggregate credit fluctuations and their implications for the real economy remains incomplete on several fronts.

Bank credit is a significant source of financing for the majority of businesses. One particularly important aspect that has been extensively studied at the micro level, yet overlooked in macro, has to do with bank-firm credit relationships. Indeed, a vast theoretical and empirical literature has long highlighted the role of these relationships in terms of alleviating agency frictions and shaping credit supply at the lender-borrower level.[1] It also emphasized the existence of cross-sectional heterogeneity in terms of match quality and inherent relationship characteristics such as duration, which can potentially hinder banks' ability to adjust their credit supply in a frictionless way (Boualam (2018)). Conversely, the common view across most macro-finance models either simply assumes homogeneous borrowers and/or lenders, or abstracts from the long-term nature of financial contracts and any market frictions that may prevent banks from costlessly forming or severing these credit matches. These models thus downplay the value of relationships and their aggregate consequences and imply that banks can swiftly adjust the number of their borrowers in response to shocks. They also leave little room for analyzing the process of credit reallocation across bank-firm matches and its dynamics throughout the cycle.

This paper proposes a novel macro perspective on the process of credit intermediation. It aims to provide further empirical evidence on the key and distinctive roles played by both the intensive and extensive margins in shaping aggregate credit fluctuations. Here, we attempt to look behind such fluctuations in order to address first-order questions such as: (i) When aggregate bank credit declines by 5%, is it because the average loan size (i.e., intensive margin) drops by 5%, or is it because 5% of bank-firm matches (i.e., extensive margin) are destroyed? (ii) Does the origin of aggregate credit fluctuations matter? (iii) Do monetary policy shocks impact these margins differently?

To our knowledge, we are the first to show that banks actively adjust both the number *and* the intensity of their relationships, in response to macroeconomic shocks, and that both of these margins represent a significant source of variation in bank lending. These adjustments are somewhat analogous to the ways in which firms constantly adjust both quantity of hours worked and employment, or their capacity utilization and new capital investment.[2] This view may sound intuitive, yet — and surprisingly — a thorough analysis of the dynamics of these margins and their macroeconomic implications remains limited, if not completely absent. Furthermore, we not only establish the quantitative importance of these margins, but we also argue that they are subject to prominently different aggregate

[1] See Boot (2000) and Degryse et al. (2009) for a survey of earlier work.

[2] To some extent, our analysis of credit markets follows in the footsteps of Lilien and Hall (1986), who first decomposed the fluctuations in total hours worked into changes in employment and changes in hours worked per employed worker.

behaviors. Thus, disentangling the effects associated with each margin can prove informative about the economic mechanisms at play and the role of credit reallocation, and ultimately yield relevant policy implications.

To shed light on this process, we leverage a key source of information, the French Credit Register, which covers the commercial loan market in France, and is maintained by Banque de France. The data contain granular and nearly exhaustive records of bank-firm matches and corresponding credit exposures over the period 1998-2018. To study the properties of credit relationship flows, we develop an empirical methodology akin to the one pioneered by Davis and Haltiwanger (1992) for labor flows. Our methodology takes into consideration specific characteristics associated with credit market structure and available data. For example, we track data entries for each bank-firm match to determine the time of creation and inferred time of destruction in order to construct the associated gross credit relationship flows. We also account for cross-sectional heterogeneity and the nature of financial contracts through key attributes such as loan size, credit type and maturity, and relationship duration.

Understanding the implications of bank-firm credit relationships is a natural undertaking. However, a dearth of empirical evidence documenting their macro-level properties exists due to the paucity of extensive micro datasets over a sufficiently long period of time. In fact, earlier studies such as Dell'Ariccia and Garibaldi (2005) relied on bank-level call report data. Thus, they cannot identify the involved borrowers and can observe net intensive flows only at the bank level. As a consequence, these studies cannot disentangle extensive from intensive margins, nor precisely capture the underlying magnitude and properties of credit reallocation. Instead, we advance here a novel approach to exploit information available in credit registers, which is typically used in micro settings, to uncover new aggregate findings.

Our research establishes the following stylized facts about the extensive and intensive margins of credit:

1. Extensive and intensive margins fluctuate continuously over time. While their persistence is roughly identical, the volatility of the intensive margin is relatively higher.
2. Both margins are important at the business cycle frequency, with the extensive margin contributing about one quarter to one half of the variance in aggregate credit.
3. In the long run, the extensive margin accounts for the bulk of aggregate credit variations.

Our analysis also highlights the following features pertaining to gross credit relationship flows:

1. The creation, destruction, and reallocation of bank-firm relationships coexist through-

out the cycle.

2. Creation (inflows) and destruction (outflows) of relationships show greater volatility compared to net flows. Variations in net flows are driven mainly by inflows.
3. Outflows are more volatile for small and short-term loans and credit relationships with duration of less than one year. Inflows are more volatile for relationships with small loans or lines of credit.

Our results also highlight that credit patterns observed during or in the aftermath of an economic downturn are driven potentially by multiple combinations of extensive and intensive margin sources, suggesting that different economic mechanisms may be at play. A better understanding of the extensive/intensive origin of a credit decline and its bottlenecks can thus be relevant to the design of effective and targeted policy tools. In this context, we analyze how monetary policy gets transmitted through both extensive and intensive margin channels. We show that while the intensive margin responds immediately and strongly to monetary policy surprises, the extensive margin's response is relatively more gradual and subdued during easing regimes. We also note that the extensive margin channel is at play mainly for relatively small banks or those with flexible balance sheets.

Our empirical framework also provides us with tools to better understand the reallocation process occurring in credit markets and the channels through which bank shocks get transmitted to the real economy. In particular, we show that the excess reallocation rate of credit relationships is countercyclical, in line with the cleansing effect of recessions. In addition, yearly (excess) reallocation rates have been steadily declining over the past two decades. These results indicate the existence of factors hampering credit market fluidity and contain relevant theoretical and policy ramifications worthy of further investigation.

Literature Review. This paper aims to connect two distinct yet complementary approaches to bank credit: macroeconomic research on credit cycles and microeconomic literature on relationship banking. The literature on credit cycles has long emphasized the role of credit constraints stemming from the borrower side (starting with Bernanke and Gertler (1989) and Kiyotaki and Moore (1997)). More recently, this literature has shifted focus toward analyzing bank constraints and decisions.[3] Yet, both of these prominent strands generally abstract from long-term contracts and omit frictions stemming from market structure, and can hardly relate the role of bank or firm heterogeneity to the process of credit reallocation. At the same time, the banking literature has largely demonstrated the importance of relationships in shaping credit, albeit theoretical and empirical studies therein have

[3] See, for example, Corbae and D'Erasmo (2019) and Begenau and Landvoigt (2018).

focused mostly on the micro level with limited macroeconomic implications.

The paper's contribution resides in providing empirical foundations and assessing the macroeconomic relevance of a complementary line of research that puts forward a flow-driven approach to credit markets. This approach posits the essential role of bank-firm match dynamics in aggregate credit and has been introduced theoretically in Den Haan et al. (2003) and Becsi et al. (2005). More recently, Boualam (2018) builds a general equilibrium model featuring frictional credit markets and long-term contracts and argues that the destruction of bank-firm relationships during crises can significantly slow down recoveries. In a related micro approach, Mazet-Sonilhac (2020) empirically investigates how the reduction in search frictions, driven by the introduction of broadband internet, impacts bank-firm matching and aggregate credit flows. From a methodological standpoint, our approach is closely related to the one commonly used to examine job flows, and specifically to earlier studies conducted by Steven Davis and John Haltiwanger and partly summarized in Davis and Haltiwanger (1999).

Our work is also part of a nascent literature on credit flows and reallocation, which includes Dell'Ariccia and Garibaldi (2005), Herrera et al. (2011), Craig and Haubrich (2013), and Contessi and Francis (2013). One closely related paper to ours is by Dell'Ariccia and Garibaldi (2005), who use bank-level information to track credit flows along the intensive margin. We argue that the use of bank-level data, while informative about flows "on the surface," in fact masks the extent of credit reallocation and cannot provide information about the dynamics of bank-firm relationships.[4] In the same vein, Herrera et al. (2011) work with firm-level data to measure inter-firm credit reallocation. Although that paper provides a valuable first step toward our understanding of credit reallocation, its focus is not on bank credit, but rather on a broad definition encompassing all forms except trade credit. In addition, the Compustat data used in that analysis cannot fully capture the extent of reallocation across borrowers and lenders, nor account for relatively small firms. In contrast, our paper is the first to use loan-level data to carefully establish patterns and stylized facts about gross credit relationship flows in order to distinguish extensive and intensive margin effects and uncover the underlying reallocation dynamics. Our unit of observation is the bank-firm match, which allows us to precisely measure credit reallocation at the loan level. This level is key because inter-bank reallocation or inter-firm credit reallocation measures tend to significantly underestimate the magnitudes of the underlying gross flows.[5]

[4]For example, banks may well be reallocating credit across their borrowers even though their net credit growth is zero.

[5]A contemporaneous research work by Cuciniello and Di Iasio (2020) also investigates a credit decomposition into extensive/intensive margins using the Italian Credit Register data, and, in line with our results,

The paper is also broadly related to the literature quantifying the sources of aggregate credit fluctuations and the transmission of shocks stemming from either borrowers or lenders.[6] Among others, three recent papers are connected to ours. Jiménez et al. (2014) identify the bank risk-taking channel of monetary policy through a bank-firm level analysis using Spanish loan application and credit register data. Amiti and Weinstein (2018) use matched bank-firm loan-level Japanese data with a focus on publicly listed companies to measure the importance of idiosyncratic granular bank supply shocks and their implications for credit and firm investment. Beaumont et al. (2019), with whom we share the use of the French Credit Register, suggest additional effects stemming from granular borrower shocks. Our paper differs from these in that it focuses on the role played by credit relationship flows, and analyzes the extensive/intensive margin decompositions of credit from an aggregate perspective.

Finally, while we have focused almost entirely on aggregate outcomes here, an ongoing companion paper, Boualam and Mazet-Sonilhac (2021), analyzes the data from a disaggregated perspective and delivers complementary results on cross-sectional properties and credit reallocation.

Organization. The paper proceeds as follows. We first start by outlining our empirical methodology and its conceptual foundations in Section 2. We then present our core results for gross relationship flows (Section 3), and aggregate credit decompositions (Section 4). Section 5 explores the properties of credit during crises and recovery and Section 6 investigates macroeconomic implications. Section 7 discusses other relevant applications and extensions and Section 8 concludes.

2. Empirical Methodology

The central objective of this paper is to shed light on the importance of the extensive margin of credit, and to document aggregate patterns and cyclical properties of gross credit relationship flows along with their intensive margin counterpart. This section first introduces conceptual foundations behind our measurements and then discusses the data and lays out our empirical methodology.

confirms the prominent role played by the extensive margin throughout the business cycle.

[6] See, for example, Hubbard et al. (2002), Khwaja and Mian (2008), Chodorow-Reich (2014), and Greenstone et al. (2020).

2.1. *Conceptual Foundations: a Flow Approach to Credit Markets*

We focus on the aggregate implications behind the dynamics of bank-firm relationships. It is thus essential to disentangle the extensive and intensive margins of credit. We start with a simple credit market identity, which states that total aggregate credit supplied by banks, C_t, is the product of the number of credit relationships (i.e., extensive margin), N_t, which we refer to as *relationship capital*, and the average credit exposure per relationship (i.e., intensive margin), $\bar{c}_t$:

$$C_t \;\; = \;\; N_t \times \bar{c}_t. \tag{3.1}$$

This decomposition presupposes that all firms are ex-ante identical and that credit relationships are all homogeneous. Although this approach potentially masks compositional effects, it has the merits of being straightforward and easy to interpret and measure. In that sense, the underlying changes in the extensive and intensive margins shape the dynamics of aggregate credit. Furthermore, we can write down the dynamics of relationship capital as follows:

$$N_{t+1} \;\; = \;\; N_t + \mathbb{C}_{t+1} - \mathbb{D}_{t+1}, \tag{3.2}$$

where $\mathbb{C}_{t+1}$ and $\mathbb{D}_{t+1}$ represent creation and destruction flows materialized between times t and $t+1$, respectively. These creation and destruction flows can take multiple forms within credit markets. Figure 3.1 represents them conceptually from the firms' perspective. We consider that firms can be in one of two states: (i) funded, or (ii) unfunded. Creation flows can thus represent the formation of a bank-firm relationship (as the unfunded firms become funded), but also situations where already funded firms switch banks or accumulate multiple banking relationships. In a similar vein, destruction flows represent the severance of credit relationships. These destructions can be viewed as "internal" to the credit market, as is the case for firms transitioning from funded to unfunded states, switching banks, or separating from part of their established banking relationships. These flows can also be "external," whereby the bank-firm match destruction is due to permanent firm exit or default.

In the spirit of the labor market literature, we rely on search theory insights to provide the conceptual grounds underlying our measurements of gross credit relationship flows. We also follow general insights from Boualam (2018), who posits that credit markets are subject to imperfections, akin to search frictions, and that this may lead to a form of asymmetric adjustment costs in bank relationship capital. In addition, the total credit intermediated is

a function of the number of relationships but also of their intensity and composition. As search is costly, banks and firms spend time looking for matches. Frictions affecting the matching, severance, and reallocation of bank-firm pairs acts as a form of credit adjustment cost. We view credit relationship creation and destruction as inherent to a large process of adjustment and reallocation of capital across banks and firms. One key insight is that these adjustments are time-varying, with potentially asymmetric costs associated with creation and destruction and extensive and intensive margins. These costs may also depend on bank, firm, and credit relationship characteristics. Thus, we exploit concepts laid out by this flow-driven approach, and construct novel measures that shed light on the structure of credit markets and bank-firm relationships.

2.2. Definitions and Measurement

Next we propose new measures for credit relationship flows, credit exposure, and relationship intensity and lay down their underlying assumptions and interpretations. Our definitions rely on the conceptual foundations above and follow in the footsteps of earlier studies of gross job flows (Davis et al. (1998)).

Credit Relationship (CR) Flows

We start with the definition of creation and destruction flows and credit relationships.[7]

Definition 1.

- ***Credit Relationship Creation (inflow).*** *First occurrence of a bank-firm match with strictly positive credit exposure at time t, assuming no previous match over the preceding 4 quarters, i.e., between t-4 and t-1.*
- ***Credit Relationship Destruction (outflow).*** *Last occurrence of a bank-firm match, assuming no further match for at least the next 4 quarters, i.e., between t+1 and t+4.*[8]
- ***Credit Relationship.*** *Existing bank-firm match at time t, whereby t lies within the creation and destruction dates.*

Our definitions put forward the theoretical construct of a credit relationship.[9] Figure 3.2

[7] An earlier version of these flow definitions was proposed in Boualam (2018).

[8] Our choice of a 4-quarter gap is informed by existing data and the fact that most "recalls" occur within the first year (for example, we find that only 10.5% of severed relationships are recreated 5 to 8 quarters later). Our results are qualitatively unchanged if we allow for the number of quarters to be 8 and 12, as shown in the Online Appendix.

[9] Throughout the paper, we refer to credit relationships and bank-firm matches interchangeably.

depicts possible configurations in the data and our measurement choices. We assume that borrower information is not lost immediately upon the expiration of a given credit facility, but only after a relatively long interaction-free period of several quarters. This approach helps to account for the fact that banks and firms may engage in lengthy negotiations before closing a loan deal, and also adjusts for cases where credit exposures temporarily decline below the mandatory reporting threshold and for potential reporting gaps in the data. Note that while the last occurrence of a bank-firm match may be at quarter t, the destruction of such match in fact happens some time between quarter t and $t+1$, and is thus accounted for at time $t+1$. Note also that our definitions do not preclude situations where banks and a firms engage in several credit relationship creation and destruction rounds throughout the sample.

We can then tabulate gross credit relationship flows, i.e., creation flows, $\mathbb{C}_t$, and destruction flows, $\mathbb{D}_t$, based on the sum of all bank-firm relationships that are either created or destroyed between times $t-1$ and t. In the same vein, we can define, at time t, the net credit relationship flows, $\mathbb{N}_{t+1}$, as the difference between inflows and outflows; the reallocation flows, $\mathbb{R}_t$, as the sum of inflows and outflows; and excess reallocation flows, $\mathbb{X}_t$, as the sum of inflows and outflows minus the absolute value of net flows:

$$
\begin{aligned}
\mathbb{N}_t &= \mathbb{C}_t - \mathbb{D}_t \\
\mathbb{R}_t &= \mathbb{C}_t + \mathbb{D}_t \\
\mathbb{X}_t &= \mathbb{C}_t + \mathbb{D}_t - |\mathbb{N}_t|.
\end{aligned}
$$

In this context, excess reallocation for credit relationships measures the extent of reallocation in excess of that needed to generate the corresponding net changes in total credit relationships. For example, simultaneous creation and destruction flows on the order of 10% do not impact the stock of credit relationships in the economy, yet imply a large level of credit reshuffling across firms and banks and an excess reallocation of 20%. We can eventually compute the corresponding flow rates (denoted with lowercase characters), by dividing the measure of flows experienced between times $t-1$ and t, by the relationship capital stock at time $t-1$, N_{t-1}.

Relationship Intensity: Credit Exposure, Type, Maturity, and Duration

Besides bank and firm characteristics, the credit relationship itself can be characterized along several dimensions that define the intensity of a match. We consider here three measures, namely (i) credit exposure, (ii) type/maturity, and (iii) duration.

Credit Exposure, Type and Maturity. We start by defining the credit exposure of a bank to a given borrowing firm, as the sum of withdrawn and undrawn credit, in addition to bank credit guarantees.[10] We further decompose the withdrawn component by maturity (i.e., short-term and long-term) and other, less common forms of credit (e.g., credit leasing, securitized debt, overdrafts limits).

Definition 2.

- ***On-Balance-sheet Credit.*** *Accounts for long-term (>1 year) and short-term (<1 year) credit.*[11]
- ***Off-Balance-sheet Credit.*** *Accounts for lines of credit and credit guarantees.*[12]
- ***Credit Exposure.*** *Sum of on-balance-sheet credit and off-balance-sheet credit.*

We characterize the intensity of the credit relationship based on the nature and maturity of credit involved. We define (i) the share of on-balance-sheet credit as a ratio over total credit exposure, and (ii) the share of long-term credit as a ratio over on-balance-sheet credit. These measures reflect the level of commitment from the bank's perspective. A credit relationship is less binding when consisting only of short-term or off-balance-sheet credit that banks can swiftly reduce following an adverse shock.

Relationship Duration. Next we consider relationship duration as another measure of the intensity of a bank-firm match. Indeed, the repeated interaction between borrowers and lenders can help gradually alleviate agency and informational frictions and eventually lead to higher credit supply over time.

Definition 3.

- ***Credit Relationship Duration.*** *The duration $d_{ij,t}$ of an ongoing credit relationship between bank i and firm j corresponds to the number of quarters between time t and its creation date.*

[10] A credit guarantee covers a debtor's liabilities in case of delinquency. It enables the borrower to contract third-party liabilities by transferring counterparty risk to the bank, thereby creating an implicit credit exposure.

[11] This definition also accounts for medium- and long-term leasing and factoring, but these categories are omitted from our calculations as they represent less than 1% of on-balance-sheet credit.

[12] This definition also accounts for securitized loans. We omit this category from our calculations as it is represents a negligible share of off-balance-sheet credit.

2.3. The French Credit Register

Our analysis essentially relies on the French Credit Register, referred to as *Service Central des Risques* (henceforth SCR). This is a monthly database that contains bank credit exposures to borrowing firms over the period 1998-2018. This is the most comprehensive commercial credit dataset maintained by Banque de France and is used to monitor overall credit supply and risk exposures of domestic banks. The data are generated from detailed mandatory reports filed by all credit institutions (classified through a unique *Code Interbancaire* (CIB) identifier) and which list any credit commitment or risk exposure to any borrowing firm (as defined by a legal unit and referenced by a unique national identification number, SIREN). Reports encompass the funds made available or drawn credits; banks' credit line and guarantee commitments; in addition to leasing, factoring, and securitized loans.

Reporting financial intermediaries account for all resident credit institutions, and investment firms. Thus, the dataset provides an extensive account of existing bank-firm linkages, provided that the credit exposure is above the legal nominal reporting threshold of EUR 75,000 for the period 1998-2005 or EUR 25,000 from 2006 onward.

Data Construction. Our sample excludes firms headquartered outside Metropolitan France, self-employed entrepreneurs, and certain types of entities such as nonprofit organizations.[13] It also omits observations related to public credit institutions, non-traditional banking groups, and non-credit intermediaries, which may have different objectives compared to more standard banks.[14] We also exclude very small institutions with credit exposures averaging less than EUR 1 million quarterly.

We choose to work at the quarterly frequency given our analysis objective and the considerable size of available data. We also construct bank-firm relationships at the bank level (instead of branch or banking group levels). We deflate all credit variables using the GDP deflator for France.[15] We similarly define the reporting threshold in real terms. Furthermore, we focus on bank-firm pairs using the inflation-adjusted threshold (corresponding to EUR 75,000 in 1998) throughout the sample period in order to make sure that our analysis remains consistent over time despite the 2006 change in reporting.[16]

[13] The Online Appendix provides additional details related to data filters and variable construction.

[14] These include Caisse des Depots et Consignations, Oseo, and Banque de Developpement des PME, which later became Banque Publique d'Investissement (BPI) in 2015. Credit supplied through public banks accounts for about 15% of the total credit over the sample period.

[15] All credit variables are reported in terms of 1998 EUR based on the GDP implicit price deflator in France constructed by the OECD and retrieved from FRED (FRAGDPDEFQISMEI).

[16] The reporting threshold is fixed to EUR 75,000 at the beginning of the sample period and is adjusted over time. Given that inflation remains positive overall throughout the sample period, this means that we

Our cleaned baseline dataset contains about 27 million bank-firm-quarter observations over the period 1998Q1-2018Q4, involving 715 unique banks (447 banks per quarter on average) and 940,554 unique firms (256,271 firms per quarter on average). Figures 3.3(a) and 3.3(b) report the evolution of the number of banks and firms and that of total credit and relationship capital, respectively. While the banking sector experienced intense consolidation over the sample period with the number of banks declining by a third, the number of firms relying on bank credit had almost doubled during that time.

Comparison with Aggregate Flow of Funds Data. Our final sample covers about 61% of total bank credit to non-financial companies as reported in the balance of payments available through the Flow of Funds data. The two time series exhibit similar aggregate patterns overall, with correlations of 0.99 for total credit and 0.98 for long-term credit.

2.4. Issues and Adjustments

Our data and empirical methodology are subject to certain limitations and other standard issues, which may tend to affect the level of relationship flows. These include (i) seasonality, (ii) bank and firm consolidations, and (iii) variations in the reporting threshold. We attempt to correct for these limitations and discuss them in detail in this section.[17]

Seasonality. The flow data exhibit strong seasonality patterns with higher creation flows in quarter 1 and higher destruction flows in quarter 4. We use the standard X-13 procedure to generate seasonally adjusted time series. Furthermore, and although such issues appear to be negligible, we smooth the data using a centered moving average (-1,1) to control for potentially mistimed reports of credit exposures.[18]

Bank Consolidation. The French banking sector has undergone several rounds of consolidations throughout our sample period. As shown in Figure 3.3(a), the number of banks declined from 547 to 342 from 1998 through 2017. This is due almost exclusively to mergers and acquisitions, as bank entry and exit events were negligible during this period. Bank consolidations may lead to spurious inflows and outflows when the acquiring bank starts

omit a small fraction of bank-firm relationships present in the data but that exhibit below-threshold credit exposures.

[17] We relegate other minor issues such as those related to the change in the reporting of some categories (e.g., off-balance-sheet credit) and the classification of non-performing credit and ensuing adjustments to the Online Appendix.

[18] For example, a loan deal that closes on December 31 might not be officially reported until the following quarter. Similarly, a relationship that gets terminated on January 1 might not be accounted for until the next quarter.

reporting the credit relationships originally attributed to the acquired bank. For example, consider a merger where Bank A acquires Bank B between times $t-1$ and t. At time $t-1$, outflows stemming from bank B, $O_{B,t-1}$, are overestimated given that Bank B stops reporting. In a similar vein, inflows tabulated from Bank A at time t are overestimated by the same amount, $O_{B,t-1}$, as transferred relationships are treated as if they were newly formed.

We remedy to this issue following the methodology in Dell'Ariccia and Garibaldi (2005) and using the list of bank merger events maintained by the French Supervision and Prudential Authority (ACPR). This list accounts for all banking M&A activity involving banks located in France over the period 1995 through 2016. We thus proceed by (i) setting acquired Bank B's outflows to zero at $t-1$ ($O_{B,t-1} = 0$) and (ii) reducing the time t inflows associated with acquiring Bank A by $O_{B,t-1}$. We omit mergers involving banks with missing identifiers, which account for less than 3% of M&A events. We also complement these adjustments by manually checking the database and accounting for more complex situations, such as the consolidation of Caisse d'Epargne and Banque Populaire.[19]

Firm Consolidation. M&A activity at the firm level could also generate spurious flows. For example, when Firm A absorbs Firm B that is linked to Bank C, our measurement definitions would record the simultaneous destruction of the B-C match and the creation of a new A-C match instead. While we cannot adjust directly for these flows in the absence of an exhaustive corporate M&A database for France, we can show that the M&A-induced flows represent a negligible fraction of our tabulated measures. We estimate the existence of fewer than 40,000 instances of firm consolidations in France over the period 1999-2018 through Bureau Van Dijk's Zephyr database, the most comprehensive dataset available to us. This number represents less than 0.15% of bank-firm credit relationships and about 2% of total gross flows over the sample period. Furthermore, when considering only the subsample of larger firms reported in FIBEN (i.e., the universe of firms for which balance-sheet information is collected by Banque de France), we estimate that a conservative upper bound for the share of M&A-induced flows is around 5 to 6%. Finally, we note that our analysis is immune from other types of activity leading to ownership or name changes for standalone companies, given that their legal identifier (SIREN) is unique and remains constant irrespective of ownership or other legal adjustments.

Reporting Threshold. Given that we consider only those bank-firm relationships for which the total credit exposure exceeds EUR 75,000, we check that the flows of relationship

[19]This case corresponds to the absorption of one banking subsidiary by multiple acquiring banks. Here, we correct for this merger through a uniform adjustment of inflows across all acquirers.

creation and destruction are not driven simply by threshold-crossing increases or declines in credit, which can mechanically generate a positive correlation between extensive and intensive margins.

While we cannot fully rule out this possibility, we carefully address it and estimate its extent through the following tests. First, our definition of creation and destruction flows is conservative, as it accounts for an effective relationship separation only if a bank-firm match has been inactive (i.e., absent from the SCR database) for four consecutive quarters. That way, temporary declines in credit, below the reporting thresholds, do not generate spurious episodes of relationship destruction followed by creation. Second, we re-run our analysis based on a EUR 25,000 reporting threshold over the period 2006-2018 and show that the obtained patterns are qualitatively and quantitatively in line with our benchmark results. Third, we can trace back a large fraction of relationship creation and confirm that a vast majority is due to either new entrant firms (based on their creation dates obtained from the SIRENE database) or bank switches. Similarly, a large fraction of relationship destruction is due to defaulting or exiting firms in addition to switches. Fourth, we show that the average credit supplied at the time of creation or destruction of credit relationships hovers around EUR 500,000, about seven to eight times higher than the reporting thresholds, which further mitigates the extent of any related bias. The Online Appendix reports robustness tests associated with the reporting threshold.

2.5. Summary Statistics and Aggregate Time Series

Table 3.1 reports summary statistics for key variables pertaining to banks, firms, and credit relationships. The average bank has 802 distinct borrowing firms with an average credit exposure of EUR 1.03 million each (based on 1998 EUR). Furthermore, this exposure consists of about EUR 413 thousand in long-term debt, EUR 214 thousand in short-term debt, and EUR 413 thousand in undrawn credit lines.

The average number of banking relationships exhibited a slight decline from around 1.45 to 1.32, with the fraction of firms engaged in a single relationship hovering around 80%. On the other hand, banks have grown bigger, and serviced about two-and-a-half times more firms in 2016 relative to 1999. Perhaps more surprising is that the average credit exposure per firm remained relatively stable in real terms throughout the sample period, around EUR one million.

The composition of debt has shifted, however, toward long-term credit and credit lines at the expense of short-term credit, as seen in Panel (a) of Figure 3.4. We see that the percentage

share of long-term credit (over total credit exposure) surprisingly increases while the share of short-term credit declines during crises, which potentially reduces banks' ability to adjust their credit exposure. This finding is consistent with firms mostly withdrawing from their long-term, pre-committed credit lines, in line with U.S. evidence (Ivashina and Scharfstein (2010)).

Banks and firms engage in relatively long-term relationships. We estimate the average duration of a match to be on the order of 15 quarters, a bit shorter than four years.[20] Tabulating the weighted average relationship duration in the economy may be subject to some biases that could lead to spuriously large (resulting from a small subset of relationships with extremely long duration) or low numbers. We therefore choose instead to track relationships classified as those with durations of below and above two years, in order to have a better sense of how the distribution of relationship durations evolves over time. Panel (b) in Figure 3.4 shows the steadily decline the fraction of relationships with duration below two years, consistent with a lower rate of bank-firm destruction.[21]

Credit exposures associated with newly created (destroyed) relationships, account for 57% (45%) of that of incumbents, which corresponds to about EUR 570 (450) thousand, well above the reporting threshold. Furthermore, the average credit amount supplied to newly created relationships (or previously supplied to exiting firms) is procyclical, suggesting that the sub-extensive margin may play an important role in aggregate credit fluctuations. We will get back to this point in Section 4.

3. Properties of Credit Relationship Flows

We now analyze the properties of credit relationship flows. Here, we show that the processes of creation, destruction, and reallocation of credit relationships are (i) significantly large, (ii) volatile, (iii) asymmetric, and (iv) inherent to credit markets at all times.

3.1. Aggregate Patterns

Figure 3.5(a) exhibits the aggregate patterns for flows. Gross credit relationship flows are inherent to credit markets and exhibit relatively large magnitudes and volatilities. They are also quite large relative to the underlying net flows. Our results suggest that about 1 in 14

[20] This is in fact a lower-bound estimate, as we assign a duration of 0 to all bank-firm matches existing in 1998Q1, the starting date of the sample. We also do not count quarters in which the bank-firm match may be missing from the SCR when its credit exposure level drops below the reporting threshold.

[21] We select this threshold because the data show a distinct behavior for very young relationships relative to the rest. The two-year threshold is also chosen mainly so as to keep the longest possible time series.

credit relationships is created and 1 in 16 is destroyed on a quarterly basis. On average, 23,407 positive flows (6.94%) and 21,497 negative flows (6.32%) combine to generate 1,910 net flows (0.62%) per quarter. As a result, the excess reallocation rate is on the order of 12.51% per quarter. Moreover, these gross flows appear to closely track each other throughout the sample period. This observation further illustrates that the substantial process of credit reshuffling across financial institutions and borrowers is continuously reshaping credit markets. We also note that both gross flows exhibit downward trends, with quarterly flow rates of relationship creation and destruction declining from about 8.6% to 6% and from 6.8% to 5.8%, respectively, over the sample period.

3.2. Cyclical Properties

We examine the cyclical properties of gross credit relationship flows and characterize the magnitude of their fluctuations. We detrend all flow rates using the Hodrick-Prescott (HP) filter with a smoothing parameter of 1600. Figure 3.5(b) presents the corresponding cyclical deviations from the HP trend. We also tabulate the volatility, autocorrelation, and correlation with respect to log-growth of GDP, aggregate credit, and relationship capital, for each variable of interest. Table 3.2 formalizes these results.

First, we establish that creation flows (measured in levels or rates) of credit relationships are two to three times as volatile as their destruction counterpart. The standard deviation of creation flows is 0.044 for levels (0.0027 for rates), while the volatility of destruction flows is 0.026 for levels (0.0013 for rates). Second, and maybe unsurprisingly, rates of creation flows are positively correlated with growth rate of GDP (0.44), aggregate credit (0.47), and relationship capital (0.64). On the other hand, rates of destruction flows exhibit only a moderately negative correlation with GDP (-0.084), aggregate credit (-0.14), and relationship capital (-0.26). We confirm results when looking at the levels of creation and destruction flows. Third, we establish that the variations of net flows are relatively large, with a volatility of 0.051. Indeed, the procyclical nature of inflows and the countercyclical nature of outflows combine to generate large movements in net flows. Fourth, we can measure the relative contribution of each component toward the overall variance of (detrended) net flows using the following decomposition:

$$1 \quad = \quad \underbrace{\frac{\text{cov}(c_t, n_t)}{\text{var}(n_t)}}_{\beta_{pos}} + \underbrace{\frac{\text{cov}(-d_t, n_t)}{\text{var}(n_t)}}_{\beta_{neg}}, \tag{3.3}$$

and show that positive flows account for about 84% of the variation in net relationship flows

while negative flows account for only about 16%.

This result is robust across sample periods. Taking the periods pre- and post- 2008 separately, we observe that the creation margin drives net flows, while the destruction rates remain relatively stable. Furthermore, we also highlight a regime shift taking place around the financial crisis, with negative flow rates declining from their 6.5-7 percent range, pre-crisis, to a 5.5-6 percent range.[22]

Overall, and from an aggregate perspective, this finding suggests that the critical adjustment variable for relationship capital is along the creation margin, as banks may have limited control over destruction flows. This result confirms earlier findings obtained in Boualam (2018) for the U.S., based on Dealscan data, but disagrees with that reported in Dell'Ariccia and Garibaldi (2005) and Herrera et al. (2011), who use bank-level and loan-level data, respectively.[23,24]

3.3. What Drives the Creation and Destruction of Credit Relationships?

Figure 3.6 presents the decomposition of creation and destruction of credit matches as follows:

- Creation flows: (i) bank switches and multi-bank firms experiencing a relationship gain ("positive reallocation") and (ii) new firm entry (with credit exposure above the threshold).
- Destruction flows: (i) bank switches and multi-bank firms experiencing a relationship loss ("negative reallocation"), (ii) firm default, and (iii) firm exit (excluding default) or with a loan below the reporting threshold closed.

A bank switch is defined as the simultaneous move from one lender to another, which corresponds to the destruction of the original relationship and the creation of a new one within a four-quarter interval. We define multi-bank firm relationship gains (losses) as the incremental addition (drop) of a credit relationship induced by firms with a (multiple) pre-existing relationship(s). Here, we also choose to report bank switches along with relationship gains/losses, as some firms appear to switch from one lender to another at a gradual pace, generating transitory periods where they are formally associated with two banks.

[22] The relative importance of gross flows is also visible in the scatter plot in Figure 3.25 in the Online Appendix.

[23] Arguably, several differences across our samples may explain this discrepancy. Among others, these studies focus on U.S. data and different sample periods. Equally important, they use data aggregated at the bank or firm levels instead of working at the credit relationship level, as we do.

[24] Interestingly, this result is also different from labor market studies, which suggest that job destruction rates are more volatile and relatively more important for net labor flow fluctuations (Davis and Haltiwanger (1992)).

Overall, we show that about two thirds of creation flows are due to new entrants, while one third is due to incumbent firms switching to or matching with additional lenders. On the other hand, we report that destruction flows are due to bank switches and multi-bank firms experiencing a relationship loss (about 40%), firm exit (about 40%), and firm default (20%). These contributions appear to be relatively stable across the sample period and each component generally inherits the cyclical properties of the underlying gross flow. However, this picture looks slightly different for net flows, which appear to be explained mostly by the spread between entry and exit flows. Furthermore, despite the substantial volume of their gross flows, incumbent borrowers who are either switching or adding/dropping credit relationships have net flows comprising only about one fourth the volume of those exhibited by entering and exiting firms.

How Important Are Firm Entry and Exit? We complement our analysis by using the SIRENE database to help us determine the dates of firm entry (i.e., the firm's legal incorporation date) into the French economy. We then tabulate the ratio of first-time borrowers (i.e., firms appearing in the SCR for the first time) over newly created firms (entrants) within the same quarter.[25] While the flows of entrant firms and those that obtain credit for the first time are highly correlated (84%), their ratio exhibits stark dynamics. As shown in Figure 3.7, the share of first-time borrowers over entrants presents a downward secular trend (declining from about 26% to 20%) and procyclical patterns, suggesting that newly created firms have harder time getting credit during crisis periods.[26]

3.4. Cross-sectional Decomposition at the Relationship Level

We further uncover the determinants behind these fluctuations by analyzing the dynamics of credit relationships as a function of their key characteristics, namely (i) credit exposure, (ii) credit type and maturity, and (iii) duration. Figure 3.8 shows the times series associated with gross flows and Table 3.3 presents the results pertaining to their cyclical properties.

Decomposition by Relationship Credit Exposure. We specify fixed dollar thresholds at 250 thousand, 500 thousand, and 1 million EUR throughout the sample, and classify bank-firm matches into small, medium, or large credit size categories on a quarterly basis.[27]

[25] Unfortunately, there is no connecting table between firm identifiers in SIRENE and SCR datasets and thus we cannot track the outcome of each individual entrant firm.

[26] While we cannot completely rule out the possibility of a significant procyclical shift in credit exposure to new entrants, our results remain unchanged even when considering the lower reporting threshold in the second half of the sample.

[27] Adjusting the size classification thresholds for inflation does not qualitatively alter the results.

Credit relationships classified by credit exposure into (i) small (below 0.25 million Euro), (ii) medium (0.25 to 0.5 million Euro), (iii) large (0.5 to 1 million Euro), and (iv) very large (above 1 million Euro) account for about 61%, 19%, 10%, and 10% of total relationships, respectively. We show that gross and net flows associated with small loans exhibit a larger volatility relative to the largest ones. They also exhibit the largest decline in inflows during downturns. This suggests that banks consider their smaller relationships as a key variable of adjustment throughout the cycle and reiterates the additional vulnerability of small borrowers during crises. Thus, small loans have a significant impact on aggregate credit fluctuations given their relative importance in terms of share in aggregate credit relationships and lending volume.

Decomposition by Credit Type and Maturity. We decompose relationships by credit type and maturity as follows: (i) credit line, (ii) short-term, (iii) long-term, and (iv) short & long-term. Here, we will classify all credit lines within one category, given the limited availability of information about their maturity.[28] We observe that credit relationships based solely on short-term credit or credit lines experience significantly larger gross flows than relationships involving long-term credit. This suggests that these credit types offer more adjustment/reallocation flexibility to banks, as they may be cheaper to originate and/or less costly to break up. This interpretation is further supported by evidence that these two categories are subject to large increases in outflows across all four crises, while long-term credit relationships flows remain relatively stable.

Decomposition by Relationship Duration. We examine the effect of relationship duration on outflows. We classify relationships into four buckets: duration of (i) below one year, (ii) between one and two, (iii) between two and five, and (iv) above five years. The average shares associated with each category are 17%, 16%, 26%, and 41%, respectively. While outflows appear to increase across all categories during most recessions (one exception being a decline in outflows for one- to two-year duration relationships in 2008), the most sensitive relationships are those that have been active for less than one year.

To summarize, mature credit relationships with large and long-term credit exposures are overall more resilient during crisis periods relative to younger and smaller ones. They are also potentially more difficult to initiate, as their inflow rates are substantially lower. These results are not necessarily surprising in light of the positive relationship between duration

[28] One could partially infer such maturity once the firm draws from the credit line and the corresponding bank later reports the corresponding amount as short- or long-term credit.

and credit size and the negative relationship between duration and separation probability.[29] As a result, the gross flow patterns we uncover may imply that different adjustment costs are at play and depend on the value of the credit relationships, as measured by their size, type, maturity, and duration. Indeed, it may be easier to sever a young or small bank-firm match if the lost value from ending such a relationship is relatively minimal. In the same vein, it is also easier to approve relatively small loans if this only marginally impacts the credit and/or counterparty risk faced by the bank.

4. How Do Banks Adjust Their Credit Supply Along Extensive and Intensive Margins?

What levers do banks use to adjust their credit supply? What is the relative importance of extensive and intensive margins in credit fluctuations? These questions are inherent to the macro-finance literature, yet they have surprisingly received very little attention. In this section, we attempt to address them by exploring two related decompositions of credit variations. We establish that accounting for the extensive margin is absolutely critical for the proper inference of bank lending behavior in the aftermath of an aggregate shock or a new policy implementation. Thus, grasping and measuring the channel through which credit market participants form or sever matches is key to understanding of credit dynamics.

4.1. Simple Credit Decomposition (Decomposition 1)

We start with the simple credit market identity described in equation (3.1) and operate a log-transformation so as to make this decomposition additive:

$$\log(C_t) \quad = \quad \log(N_t) + \log(\bar{c}_t). \tag{3.4}$$

We detrend our aggregate credit variables using the HP filter with a smoothing parameter of 1600. The standard deviation of detrended aggregate credit (in log) is 2.58%, while the standard deviation of the number of relationships is 1.14%, and that of average credit per relationship is 1.93%. The correlation of aggregate bank credit with the two latter series is 0.71 and 0.91, respectively.

[29] Using loan-level Japanese data, Nakashima and Takahashi (2018) link relationship destruction rates to bank capital constraints and show that these are more prevalent for younger matches.

4.2. Secular Trends

Figure 3.9 reports the long-run trends associated with our three variables of interest (in logs). The trends show a significant increase (about 50%) in aggregate credit (in real terms) over the past 20 years. Interestingly, this pattern has been accompanied by an almost equivalent increase in relationship capital (about 45%) and a minimal increase in average credit per match (about 5%, which corresponds to the average credit rising from about 977 thousand to 1.02 million EUR from 1999 through 2016). In fact, while the trend in the intensive margin was relatively evident in the first half of the sample (+12%), the advent of the financial crisis has led to a gradual decline over the period 2008-2017 and thus an overall negligible contribution to aggregate credit in the long run.

This relative stability of the average credit per match may suggest that firm size composition and corresponding financing needs also remained stable throughout the sample period.[30] Hence, such finding establishes that low-frequency changes in the number of relationships may be the dominant force for long-run fluctuations in aggregate credit. As a consequence, policies that aim to boost aggregate credit in the long run may be more effective when targeting structural changes that impact the matching process between borrowers and lenders and gross relationship flows in general.

4.3. Cyclical Fluctuations

We now move on to the cyclical properties. Here we start with a simple and straightforward approach based on first differences before complementing it with an analysis of log-deviations from the HP trend.

First-difference Approach

Based on the identity derived in (3.4), we can apply a first difference between time t and $t+1$ to get:

$$\Delta \log(C_{t+1}) = \log(C_{t+1}) - \log(C_t) = \Delta \log(N_{t+1}) + \Delta \log(\bar{c}_{t+1}), \tag{3.5}$$

where $\Delta X_{t+1} = X_{t+1} - X_t$. Figure 3.10(a) illustrates the evolution of the aggregate credit (log-growth) in addition to its two extensive and intensive margin components over the sample period. For the most part, the large credit declines observed during crisis periods are due to the joint effect of both margins. In addition, the extensive margin seems to exhibit a

[30] Note that only a small fraction of firms relies on more than one relationship; thus, the average credit per match is a reasonable proxy for the total credit per given firm.

"smoother" pattern and maybe a slower reaction over time, which highlights possible differences in adjustment behaviors and costs for each margin. In addition, such decomposition can also help characterize credit recoveries. For example, the nearly creditless recovery observed in 2010-2012 was due to a relatively subdued average credit per bank-firm pair, while the number of bank-firm relationships was actually growing over the same period. We further elaborate on these crisis/recovery patterns in Section 5.

Interestingly, given that the log-transformation allows for the extensive and intensive margins to be additively separable, we can write a linear decomposition of the variance of total credit flows in the spirit of Fujita and Ramey (2009) and formally quantify the contribution of each margin, as follows:

$$\mathrm{var}(\Delta\log(C_t)) \quad = \quad \mathrm{cov}(\Delta\log(N_t),\Delta\log(C_t)) + \mathrm{cov}(\Delta\log(\bar{c}_t),\Delta\log(C_t)). \tag{3.6}$$

Ultimately, we can write:

$$\begin{aligned} \beta_{Ext} &= \frac{\mathrm{cov}(\Delta\log(N_t),\Delta\log(C_t))}{\mathrm{var}(\Delta\log(C_t))}, \\ \beta_{Int} &= \frac{\mathrm{cov}(\Delta\log(\bar{c}_t),\Delta\log(C_t))}{\mathrm{var}(\Delta\log(C_t))}, \end{aligned}$$

with $\beta_{Ext} + \beta_{Int} = 1$.

Moreover, we can rewrite the change in relationship capital (in logs) in terms of flow rates:

$$\Delta\log(N_t) = \log(1+n_t) \simeq 1 + n_t = 1 + c_t - d_t. \tag{3.7}$$

Assuming that n_t is relatively small, we can derive the following first-order approximation to further decompose the contribution of the extensive margin into creation and destruction components:

$$\mathrm{var}(\log(1+n_t)) \quad \simeq \quad \mathrm{var}(n_t)\mathrm{cov}(c_t,n_t) + \mathrm{cov}(-d_t,n_t), \tag{3.8}$$

where c_t, and d_t are the credit relationship creation and destruction rates, respectively, and

thus obtain:

$$
\begin{aligned}
\beta_{Ext} &\simeq \frac{\text{cov}(\Delta \log(N_t), \Delta \log(C_t))}{\text{var}(\Delta \log(C_t))} \\
&= \underbrace{\frac{\text{cov}(c_t, \Delta \log(C_t))}{\text{var}(\Delta \log(C_t))}}_{\beta_c} + \underbrace{\frac{\text{cov}(-d_t, \Delta \log(C_t))}{\text{var}(\Delta \log(C_t))}}_{\beta_d}.
\end{aligned}
$$

Panel A in Table 3.5 reports the results of this variance decomposition. The intensive margin accounts for 73% of the total credit variation while the extensive margin accounts for the remaining 27%. In addition, positive flows drive the bulk of the variation in the extensive margin, while negative flows are much less important. With the first-difference decomposition, we also note that the contribution of negative flows reduces, rather than increases, the variance of credit, as gross flows trend downward throughout the sample period. Once we detrend flow variables, we see that negative flows have a negligible impact on the extensive margin and aggregate credit more generally.

HP Filter Approach

We complement these results by studying cyclical deviations from HP trends. Table 3.4 reports the correlation structure while Figure 3.10 illustrates the evolution of the two margins for both the first-difference and HP filter approaches.[31] As shown earlier, the relative contributions are about one quarter for the extensive margin and three quarters for the intensive margin (Panel B in Table 3.5). We highlight that a non-negligible number of quarters with minor fluctuations in aggregate credit may actually be experiencing counteracting extensive and intensive margin effects. Furthermore, the relatively low correlation between extensive and intensive margins (0.25 based on log-growth, and 0.46 for log-deviations) also suggests that each component responds differently to aggregate shocks.

4.4. Alternative Decomposition: Incumbent vs. New and Severed Credit Relationships and the Importance of the Sub-extensive Margin (Decomposition 2)

Next, we complement our first decomposition with alternative and more refined versions in order to account for heterogeneity across incumbent, new, and severed bank-firm relation-

[31] While the decomposition into extensive/intensive margins remains straightforward, disentangling the respective effects of gross relationship flows requires additional derivations and approximations, detailed in the Online Appendix.

ships, which is prevalent across our data. We define and quantify the role of the sub-extensive margin and show that it further amplifies the distinctive features of extensive and intensive margins.

In order to justify the role of the sub-extensive margin, we first show in Figure 3.11 that the average credit size of entering and exiting borrowers corresponds to about 50% and 40% of that of the average incumbent, respectively. Moreover, this credit ratio for new borrowers is volatile and procyclical, consistent with theory in Boualam (2018). Similarly, the credit ratio for severed relationships also exhibits a procyclical pattern, albeit with slightly less volatility.[32]

We denote by n^ν, and $\bar{C}^\nu$, with $\nu \in \{i, n, s\}$, the number of relationships and the average credit associated with incumbent (i), new (n), and severed (s) bank-firm relationships, and observe that we can write the total credit C_t at time t, based on (future) surviving relationships (i.e., $t+1$ incumbents), combined with credit lost from relationships severed between t and $t+1$. We can also write C_{t+1} at time $t+1$, based on the existing relationship (i.e., the same $t+1$ incumbents) combined with the credit supplied to relationships newly formed between t and $t+1$. We can then formulate the following alternative decomposition of credit flows:

$$\begin{aligned} C_t &= n^i_{t+1}\bar{C}^i_t + n^s_{t+1}\bar{C}^s_{t+1} \\ C_{t+1} &= n^i_{t+1}\bar{C}^i_{t+1} + n^n_{t+1}\bar{C}^n_{t+1}, \end{aligned}$$

and write the corresponding first-difference identity for aggregate credit:

$$\Delta C_{t+1} = C_{t+1} - C_t = n^i_{t+1}\Delta C^i_{t+1} + n^n_{t+1}\bar{C}^n_{t+1} - n^s_{t+1}\bar{C}^s_{t+1}. \tag{3.9}$$

With $\alpha^j_t = \frac{n^j_t}{n^i_t}$ and $c^j_t = \frac{\bar{C}^j_t}{\bar{C}^i_t}$ for $j \in \{n, s\}$, we can write the counterpart to equation (3.5) as:

$$\Delta \log(C_{t+1}) = \underbrace{\Delta \log(\bar{C}^i_{t+1})}_{\text{Incumbent bank-firm effect}} + \underbrace{\log(1+\alpha^n_{t+1}c^n_{t+1})}_{\text{New bank-firm effect}} - \underbrace{\log(1+\alpha^s_{t+1}c^s_{t+1})}_{\text{Severed bank-firm effect}} \tag{3.10}$$

We decompose the variance in aggregate credit, in terms of an incumbent relationship effect (intensive margin), and a new and severed relationship effects, which jointly account for the

[32]These observations are consistent with credit that is increasing, and separation probability that is decreasing with relationship duration, as we show in Figure 3.21 in the Online Appendix.

extensive margin:

$$\begin{aligned}\text{var}(\Delta\log(C_t)) &= \text{cov}(\Delta\log(\bar{C}^i_{t+1}),\Delta\log(C_t))\\ &\quad +\text{cov}(\log(1+\alpha^n_{t+1}c^n_{t+1}),\Delta\log(C_t))\\ &\quad +\text{cov}(-\log(1+\alpha^s_{t+1}c^s_{t+1}),\Delta\log(C_t)). \end{aligned} \tag{3.11}$$

More importantly, this decomposition accounts for the sub-extensive margin of credit, by allowing for time-variation in the average credit size supplied to entering or exiting firms, relative to incumbents.[33] Eventually, we can write the betas associated with each component and the final decomposition as:

$$\beta_{Incumbent} = \frac{\text{cov}(\Delta\log(\bar{C}^i_{t+1}),\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))}; \tag{3.12}$$

$$\beta_{New} = \frac{\text{cov}(\log(1+\alpha^n_{t+1}c^n_{t+1}),\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))} \simeq \frac{\text{cov}(\alpha^n_{t+1}c^n_{t+1},\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))}; \tag{3.13}$$

$$\beta_{Severed} = \frac{\text{cov}(-\log(1+\alpha^s_{t+1}c^s_{t+1}),\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))} \simeq \frac{\text{cov}(-\alpha^s_{t+1}c^s_{t+1},\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))}, \tag{3.14}$$

with $\beta_{Incumbent}+\beta_{New}+\beta_{Severed}=1$.

Figure 3.12 reports the corresponding time series, and Panel B in Table 3.5 reports the decomposition results.[34] The results for this decomposition show that the incumbent effect accounts for 54% while the extensive margin (i.e., the combination of new-bank-firm and severed-bank-firm effects) accounts for 46%. The approximative results obtained for the HP filter decomposition confirm as well the relatively balanced contribution between the two margins. Looking at the effects of new and severed relationships separately, we find that their βs are 0.62 and -0.17 (0.57 and -0.17 for the HP filter approach), respectively. The negative sign associated with $\beta_{Severed}$, while surprising, comes in part from the fact that the average credit of recently destroyed relationships is procyclical and the corresponding fluctuations actually dominate those associated with the flow component.

In summary, we note that the extensive margin contributes significantly to aggregate credit fluctuations at the business cycle frequency. In fact, through our first decomposition, we estimate that such relative contribution is on the order of one quarter. When we also take

[33] While we report in the core part of the paper the derivations for the first-difference approach, the details associated with the HP filter approach are in the Online Appendix.

[34] Section D.3 in the Online Appendix presents a third decomposition (referred to as "gross intensive credit flows"), which further disentangles the positive and negative flows within the intensive margin. Its results (for both first-difference and HP filter approaches) are reported in Table 3.8 and are overall consistent with the ones presented in this section.

into account the sub-extensive margin, and hence the extent of heterogeneity (in terms of credit size) existing between incumbent and new or severed credit relationships, we see that the relative contribution of the extensive margin is actually even larger and jumps to about one half.

5. Anatomy of a Credit Crisis and Recovery

We have so far highlighted the importance of both extensive and intensive margins in explaining fluctuations of aggregate credit. In this section, we trace out the anatomy of a credit cycle by constructing the average cyclical fluctuations of our variables around economic downturns with a focus on the second decomposition approach. The French economy experienced four recessions over the period 1998-2018 according to the OECD: (i) 2001-2003, (ii) 2008-2009, (iii) 2011-2013, and (iv) 2014-2016. Figure 3.13 shows the unconditional results while Figure 3.14 zooms into each recession period separately. The aggregate credit time series we report corresponds to the sum of the detrended extensive and intensive margins. Here, all variables of interest are normalized to 0 at the onset of a recession and their dynamics are then reported over the subsequent three years.

We see that aggregate credit gradually declines on average to about 7% three years after the quarter of recession onset. This decline is equitably due to both extensive and intensive margins, which fall by about 3.5% each. Interestingly, we show that the intensive margin is the main driver of aggregate credit decline in the short-run, while the contribution of the extensive margin is concentrated between the fifth and twelfth quarters. Both margins appear to be particularly persistent and start reverting back toward pre-crisis levels only after about three-and-a-half years.

With respect to gross credit relationship flows, we show that a typical recession is characterized by a sharp and prolonged decline in inflows that persists over the first four quarters. In addition, inflows remain subdued for a relatively long period, recovering only halfway from their pre-crisis level after seven to eight quarters. The persistently low level of credit supplied to newly formed relationships further amplifies the role of the creation margin during crisis and recovery periods. Conversely, outflows observe only a modest and short-lived increase in downturns. The magnitude of their change is roughly one-sixth that of inflows and they revert back to their pre-crisis levels in about four quarters. Furthermore, the relationships severed during downturns in general consist of smaller loans relative to normal periods, which further mitigate the impact of the credit destruction component.

When we zoom into each recession separately, we observe that the credit declines experienced

have different origins, namely: (i) the extensive margin as in the crisis of 2008-2009, (ii) a combination of extensive and intensive margins as in 2001-2003, and (iii) the intensive margin as in 2012-2013 and 2014-2016.[35] For example, in 2001-2003, when aggregate credit fell by about 12 log-points over the ten quarters following its peak, the intensive margin declined by eight log-points, while the extensive margin declined only by four log-points. Conversely, the credit decline of about five log-points observed in 2008-2009 was explained principally by the four log-point decline in the extensive margin.

6. The Extensive Margin Channel of Monetary Policy

How does monetary policy impact aggregate credit dynamics through the lens of the extensive and intensive margins? We address this question by estimating impulse responses using the local projection methods of Jordà (2005). Throughout this section, we focus on the extensive/intensive margins, as determined by our second decomposition approach, described in section 4.4, and on the effects of monetary policy surprises over a horizon h from zero to eight quarters. Our analysis is first conducted at the aggregate level. We then complement our findings with micro-level results and further investigate how bank characteristics can affect monetary policy transmission along both credit margins.

Let us denote by Y_t the dependant credit variable. The general specification we use is as follows:

$$Y_{t+h} = \alpha_h + \beta_h V_t + u_{t,h}, \tag{3.15}$$

with V_t, the treatment variable representing the instrument for exogenous variations in the monetary policy stance at time t, and the error term $u_{t,h}$. The estimated coefficients $\{\beta_h\}_{h=1..8}$ determine the cumulative impulse response path of the credit variables of interest following a change in the treatment variable. Given the nature of the credit decomposition (3.10), the dependent variable Y_{t+h} stands for (i) $\Delta \log(C_{t,t+h}) = \log(C_{t+h}) - \log(C_t)$ for aggregate credit, (ii) $\sum_{l=1}^{h} \log(1+\alpha^n_{t+l}c^n_{t+l}) - \log(1+\alpha^s_{t+l}c^s_{t+l})$ for the extensive margin, and (iii) $\Delta \log(\bar{C}^i_{t,t+h}) = \log(\bar{C}^i_{t+h}) - \log(\bar{C}^i_t)$ for the intensive margin.

[35] These results also extend to the first credit decomposition into relationship capital and average credit per relationship, as reported in the Online Appendix.

6.1. Measurement of Monetary Policy Shocks

Our analysis relies on the Eurozone monetary policy shocks constructed in Jarociński and Karadi (2020) over the period 2002-2018. These shocks are determined building on the high-frequency identification (HFI) methodology established in Gürkaynak et al. (2005). The HFI approach is particularly useful given that short-term monetary policy, as reflected in the Euro Overnight Index Average (EONIA), remained anchored at the zero lower bound for a significant portion of our sample period, and that monetary policy decisions are correlated with macroeconomic conditions. The identified monetary policy shocks combine surprise movements in the three-month to two-year EONIA swap rates. The surprises are tabulated within short intraday windows surrounding policy announcements (30 minutes) and press conferences (90 minutes) following the monetary policy meetings of the European Central Bank's (ECB's) Governing Council. The identification assumes that changes in interest rates and asset prices occurring within these time windows ought to be due solely to monetary policy news.

Furthermore, the ECB's transparent communication may convey substantial information about both monetary policy stance and economic outlook. Such confounding information can lead to counterintuitive results. As a result, Jarociński and Karadi (2020) disentangle these surprises into "purified" monetary policy and central bank information components by exploiting the differentiated reaction of stock market prices using the EURO STOXX 50 index and imposing co-movement restrictions into their VAR specification. The key idea is that positive shocks attributed to "purified" monetary policy generate a positive response to interest rates but a negative one to asset prices, while positive shocks attributed to central bank information generate a positive response for both interest rates and asset prices. Ultimately, our analysis delves extensively into credit responses to monetary policy shocks associated with the"purified" monetary policy surprises.[36] The shocks are aggregated from the ECB meeting to the quarterly frequency to be consistent with the remaining variables in the specification.

6.2. Aggregate Response

We estimate the specification (3.15) for aggregate credit in addition to its extensive and intensive margin components. Figure 3.15 reports the impulse response results based on

[36] Figure 3.29 reports these time series in the Online Appendix. We thank Peter Karadi for sharing these data with us.

the estimated $\{\beta^h\}_{h=1..8}$.[37] Panel (a) shows that a 100 basis point contractionary monetary policy shock produces a gradual decline in aggregate credit, reaching about 20% after seven quarters. Notably, our findings suggest that the bank lending channel operates through both extensive and intensive margins, albeit with different sensitivities and timing, as shown in Panels (b) and (c). The bulk of the credit variation appears to be driven by the intensive margin in the short run, with the extensive margin's response initially muted. Ultimately, the effect on the extensive margin becomes more prominent between the fourth and eighth quarters.

Finally, Panels (d) and (e) report differentiated responses for the creation and destruction components of the extensive margin. More specifically, we show that monetary policy surprises are channelled initially through the formation of new relationships and the average credit per new relationship, and only start impacting the destruction side after five to six quarters. This observation further confirms our earlier results that highlight the significant role of creation flows, given that banks may have limited ability to sever credit relationships featuring long-term loans.

One interpretation of the time lag between extensive and intensive margins is that it reflects higher adjustment costs related to search frictions or information asymmetry associated with the creation or destruction of bank-firm matches. As a result, monetary policy easing conducted in the aftermath of a downturn could appear to benefit solely incumbent borrowers through their existing credit relationships, at the expense of first-time borrowers. This may generate inefficiencies in credit allocation and impact small- and medium-size businesses as they delay entry decisions and access to credit.

Easing vs. Tightening Effects. We also test whether credit variable responses to monetary policy are subject to non-linearities conditional on the easing or tightening natures of the shocks, and run the following adjusted specification:

$$Y_{t+h} = \alpha_h + \beta_h^1 \times \delta_T + \beta_h^E V_t + \beta_h^T \delta_T \times V_t + u_{t,h}, \tag{3.16}$$

with a dummy variable, δ_T, that equals 0 for an easing shock and 1 for a tightening shock; $\{\beta_h^E\}_{h=1..8}$, the estimated response to an easing shock, and $\{\beta_h^T\}_{h=1..8}$, the estimated response

[37] We also report the results obtained for the Central Bank information component in the Online Appendix. We also provide further robustness tests. In particular, we highlight that our results are robust to (i) introduction of two lags of monetary policy surprises as controls, and (ii) using shocks constructed in Kerssenfischer (2019), with methodology that notably allows for a wider time window around ECB announcements and uses German Bund futures.

to a tightening shock.

Figure 3.16 shows the results. While the effects appear to be roughly of the same order of magnitude at first pass, only the responses to tightening shocks are in fact statistically significant. Thus, the overall effect of monetary policy shocks on total credit (-10%) is driven mainly by the negative effect implied by tightening surprises, as easing surprises generate non-significant increases in total credit.[38]

6.3. *Bank-level Response*

Next we revisit our impulse response results and analyze micro-level responses using bank-level information. The adjusted specification we use for local projections is as follows:

$$Y_{i,t+h} = \alpha_{i,h} + \beta_h V_t + u_{i,t,h}, \tag{3.17}$$

with V_t, the treatment variable representing the monetary policy shock and the error terms $u_{i,t,h}$.

We estimate the above panel regression (3.17) and show the estimated cumulative impulse responses in Figure 3.17.[39] Overall, the estimated $\{\beta^h\}_{h=1..8}$ coefficients (derived based on equal weights across banks) are estimated with tighter confidence intervals, given bank-level heterogeneity. The results are qualitatively in line with those estimated at the aggregate level, with the exception of the destruction component. At the bank level, we find that a 100 basis point contractionary monetary policy shock leads to an average decline in credit of about 38% after seven quarters. This is due overall to a 14% decline in the extensive margin combined with a 24% decline in the intensive margin. The notably different pattern for the destruction component of the extensive margin (Panel (e)) illustrates potentially distinct lending behaviors across the bank size distribution and suggest that smaller banks may respond to monetary policy tightening by severing fewer credit relationships or relationships that feature relatively small credit exposures, as opposed to larger banks.

What is the role of bank characteristics in the transmission of monetary policy shocks? We take advantage of the bank-level approach and investigate the above results by classifying banks into subgroups associated with (i) bank size, (ii) share of long-term loans, and (iii) share of off-balance-sheet credit items. Banks are reclassified on a quarterly basis in below-

[38] This result appears to be particularly strong for the period pre-2008. In fact, we show in auxiliary results that simultaneous Long-Term Refinance Operations (LTRO) announcements significantly mitigate the effect of conventional monetary policy surprises on credit.

[39] The results are qualitatively similar when including bank fixed-effects, as shown in the Online Appendix.

and above-median groups. We re-estimate our local projection specification (3.17) and report the corresponding results in Figure 3.18.

While credit variable responses appear to consistently follow similar paths across all subgroups, the magnitude of the effects can differ substantially. For example, banks that are smaller, with fewer long-term loans, or with more off-balance-sheet credit exposure appear to be twice as reactive to monetary policy changes relative to their counterparts. This effect is visible along the intensive margin across all characteristics. It is all the more stark for the extensive margin given that banks that are larger, with a large share of long-term loans, or with a small share of off-balance-sheet items, exhibit only a muted response, as opposed to a significant decline experienced by their counterparts. Thus, banks that retain sufficient flexibility in their balance sheet may be able to react more swiftly to monetary policy surprises. From a policy perspective, these banks may play a key role during credit recoveries. Indeed, monetary easing rounds in the aftermath of an economic downturn could specifically benefit incumbent borrowers of such banks, all else being equal. In a similar vein, these banks are likely to expand their loan supply along their extensive margin and offer better access to credit for first-time borrowers.

Taken together, our findings show that the extensive margin is critical in the transmission of monetary policy to commercial credit. This result goes beyond the bank lending channel literature, which typically focuses on the intensive margin. The extensive margin channel we uncover is central for small banks and those with flexible balance sheets and ample lending capacity.

7. Discussion - Credit Reallocation and Theoretical Implications

7.1. Credit Reallocation and Credit Market Fluidity

Trend. French credit markets have become much less fluid over the past two decades. As shown in Figure 3.5, credit market fluidity, which we define as total reallocation flows (i.e., the sum of creation and destruction flows), has declined from about 15.4% to 11.8% during that period, with the small credit segment and low-duration matches being particularly impacted.

Exploring the determinants of this substantial decline in credit market fluidity is beyond the scope of this paper. However, we elaborate below on potential contributing factors. First, a

lower degree of reallocation can be interpreted as a slower arrival of new credit opportunities and potentially longer credit search periods for newly created businesses. It can also be viewed as resulting from higher switching costs for incumbent borrowers (with potentially more monopoly rents extracted by banks), which can limit their ability to grow or to find a banking partner that better matches their needs.

Various government policies and recent banking developments may also be at play in the long run. These include bank consolidation, increased competition, tightened regulatory requirements, securitization, development of secondary markets, and improved creditor protection, and can also relate to innovations in lending technology. For example, easier access to more information, while lowering matching costs, might also prompt more precise screening, and thus to tightened lending standards. This could lead to longer credit search periods for firms, as banks become pickier, but could also generate a lower incidence of destruction of bank-firm pairs as match quality improves.

Ultimately, it is unclear whether such a trend is a considerable source of concern without a more refined exploration of bank-firm match quality and the reasons behind the reallocation slowdown. More specifically, while an increase in duration can add value for a healthy credit relationship, it could also be detrimental to the economy in the case of unhealthy ones. The lack of credit market fluidity could also have indirect implications for firm entry if borrowing is impeded and search periods are long, and for firm exit if capital remains allocated to low-quality borrowers for too long, consequently hampering productivity growth.

Cyclical Properties. A substantial literature has focused on exploring the sullying or cleansing effects of crises. Indeed, a procyclical reallocation is associated with a sullying effect to the extent that lower-quality matches tend to last longer during downturns. In the context of credit markets, this could materialize in the form of a firm's decreased ability to switch lenders as bank entry and competition decline in bad times. It could also reflect incentives that certain capital-constrained banks have to prolong credit to distressed borrowers (i.e., zombie lending). Conversely, the cleansing effect associated with countercyclical reallocation can emerge when bad matches (be they due to bad banks or borrowers) are severed and capital is efficiently reallocated toward higher quality and more resilient matches.

In the data, we find that credit reallocation is procyclical while credit churning (i.e., excess reallocation) exhibits mildly countercyclical dynamics. Table 3.2 shows that reallocation is positively correlated with the growth rate of GDP, credit, and relationship capital. It also shows that excess reallocation is negatively correlated with GDP (in both levels and rates). It is, however surprisingly, overall uncorrelated with total credit or relationship

capital over our sample period. This finding highlights the assumption that credit reshuffling across firms and banks is a key mechanism and a potential determinant behind aggregate economic performance. These results are in line with studies performed in other contexts, such as Herrera et al. (2011), who show that credit reallocation is procyclical (from the firm perspective), and Eisfeldt and Rampini (2006), who show that physical capital reallocation is procyclical. A deeper understanding of these effects requires a more refined cross-sectional analysis, which we pursue in Boualam and Mazet-Sonilhac (2021).

7.2. Implications for Theories of Banking and Credit

Our results call for a deeper understanding of the process of credit intermediation and tradeoffs faced by banks along both extensive as well as intensive margins. This is critical for the development of macro-finance models that appropriately capture credit dynamics.

Most macroeconomic banking models have so far focused on broad indicators such as aggregate credit and interest rates as their central variables, and have rarely tackled the underpinnings of credit intermediation and the aggregate implications of long-term financial contracts. Thus, since these models typically assume one-period same-size loans to which banks can make frictionless adjustments in response to shocks, they are unlikely to provide economic grounds for the existence of credit relationship flows across firms or banks, nor distinguish the role of bank-firm match heterogeneity. Given their quantitative importance, we argue that a successful theory of aggregate credit fluctuations should take into account both extensive as well as intensive margins, and carefully lay out the driving forces that may affect them differently.

First, several mechanisms and constraints could potentially shape economic tradeoffs between these two margins. On the one hand, banks may be interested in making loans to as many borrowers as possible so as to diversify idiosyncratic risk, learn more about their local environment, or supply credit beyond the limited demand of their existing customers. On the other hand, banks may be willing or simply constrained to focus on a small number of important relationships, when borrower acquisition and monitoring costs, or the marginal benefit of in-depth credit relationships simply outweighs diversification. In the same vein, if credit relationships turn profitable only in the long run, banks may be willing to spend extra effort to retain their incumbent borrowers, instead of creating new relationships. The severance of credit relationships could also lead to the destruction of bank-firm-specific relationship capital, which can be detrimental to both parties. This is the case, for example, when the match quality cannot be transferable due to informational frictions or other agency problems.

Other constraints can influence the extensive/intensive margin tradeoff. For example, the adjustment in bank credit is lumpy due to the very nature of bank-firm relationships and loan contracts. This is the case for banks with long-term credit exposures that cannot be reduced immediately following adverse shocks, but that may have some flexibility in partially adjusting their short-term credit and credit lines.

Second, the credit relationship flows we uncover result from firms' and banks' search, approval, and rollover decisions. This provides a novel perspective on the intermediation process in credit markets and the frictions therein. Thus, decisions related to the creation and destruction of credit relationships, but also to credit market entry and exit, may themselves be subject to time-varying costs linked to aggregate or idiosyncratic shocks and frictions hindering swift adjustments.[40]

In the spirit of arguments laid out by Rogerson and Shimer (2011) highlighting the utility of search models in labor, a search-theoretic approach can help make sense of empirical regularities and make predictions about borrower flows between unfunded and funded stages and across banks. It can also be useful in terms of the modeling of agents' decisions and can thus provide further understanding of the dynamics of aggregate variables (e.g., aggregate credit, interest rates) that are typically analyzed in the context of models abstracting from search frictions. For example, a decline in aggregate credit relationships may be due to the fact that borrowers are not entering credit markets, are not searching intensively, or are simply more picky with respect to lenders and corresponding contractual terms. On the other hand, it can also be due to the fact that banks have implemented higher lending standards resulting in higher rejection rates, or have decided to stop rolling over certain loans. Such alternative possibilities would be difficult to identify and quantify in standard banking models.

Search-and-matching frictions also provide new foundations for adjustment costs faced by banks along the extensive margin, and thus a relevant framework that can generate some dampening but can also further persistence in credit dynamics following adverse aggregate shocks. Thus, when the formation of new matches is costly and shocks are small or transitory, banks may focus on adjusting credit along the intensive margin. On the other hand, banks may ultimately downsize their relationship portfolio when subject to more severe or permanent shocks. This could lead to the severance of relationships, the loss of match-specific capital, and more persistent effects, especially when relationships are time-consuming and costly to rebuild.

[40] If credit relationships were homogeneous and banks could adjust them symmetrically and frictionlessly, then studying relationship flows may not be of the first order, but this is not the case.

8. Conclusion

Our analysis highlights the role and importance of the extensive margin in aggregate credit fluctuations. The methodology we develop for relationship flows extends that of labor research to account for the specificities of credit markets and is applicable to the study of other markets and countries with available credit register data. We view our empirical approach and dataset as a novel laboratory and as a first step toward uncovering more properties of credit relationships and their aggregate implications.

While we focus mostly on establishing stylized facts and identifying the distinctive features of extensive/intensive margins, we believe that fleshing out the potential economic mechanisms behind these dynamics can provide additional connections to the macro-finance literature, and in particular the role played by collateral and bank balance-sheet channels.

More broadly, given the key role played by the extensive margin of credit along the business cycle frequency and in the long run, this analysis raises the issue of whether banking models abstracting from such a quantitatively important dimension provide a reasonable benchmark for the study of aggregate credit fluctuations. Thus, building models that account for both margins is, in our opinion, critical when thinking about bank credit. We leave the implications of these arguments for future research.

References

Amiti, M. and Weinstein, D. (2018). How much do idiosyncratic bank shocks affect investment? evidence from matched bank-firm loan data. *Journal of Political Economy*, 126(2):525–587.

Beaumont, P., Libert, T., and Hurlin, C. (2019). Granular borrowers. Technical Report No. 3391768, Universite Paris-Dauphine Research Paper.

Becsi, Z., Li, V., and Wang, P. (2005). Heterogeneous borrowers, liquidity, and the search for credit. *Journal of Economic Dynamics and Control*, 29(8):1331–1360.

Begenau, J. and Landvoigt, T. (2018). Financial regulation in a quantitative model of the modern banking system. Technical report, Available at SSRN 2748206.

Bernanke, B. and Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *The American Economic Review*, 79(1):14–31.

Bernanke, B. S. (1983). Non-monetary effects of the financial crisis in the propagation of the great depression. Technical report, NBER Working Paper.

Boot, A. W. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, 9(1):7–25.

Boualam, Y. and Mazet-Sonilhac, C. (2021). Credit market fluidity and borrower reallocation. Technical report, Working Paper.

Boualam, Y. M. (2018). Credit markets and relationship capital. *Working Paper at University of Pennsylvania,*.

Chodorow-Reich, G. (2014). The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis. *The Quarterly Journal of Economics*, 129(1):1–59.

Cieslak, A. and Schrimpf, A. (2019). Non-monetary news in central bank communication. *Journal of International Economics*, 118:293–315.

Contessi, S. and Francis, J. (2013). U.s. commercial bank lending through 2008: Q4: new evidence from gross credit flows. *Economic Inquiry*, 51(1):428–444.

Corbae, D. and D'Erasmo, P. (2019). Capital requirements in a quantitative model of banking industry dynamics. Technical report, NBER Working Paper No. 25424.

Craig, B. and Haubrich, J. (2013). Gross loan flows. *Journal of Money, Credit and Banking*, 45(2-3):401–421.

Cuciniello, V. and Di Iasio, N. (2020). Determinants of the credit cycle: a flow analysis of the extensive margin. *Bank of Italy Temi di Discussione (Working Paper) No. 1266.*

Davis, S. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, 107(3):819–863.

Davis, S. and Haltiwanger, J. (1999). Gross job flows. *Handbook of Labor Economics*, 3:2711–2805.

Davis, S., Haltiwanger, J., Schuh, S., et al. (1998). *Job creation and destruction*, volume 1. The MIT Press.

Degryse, H., Kim, M., and Ongena, S. (2009). *Microeconometrics of Banking Methods, Applications, and Results.* Oxford University Press.

Dell'Ariccia, G. and Garibaldi, P. (2005). Gross credit flows. *The Review of Economic Studies*, 72(3):665–685.

Den Haan, W., Ramey, G., and Watson, J. (2003). Liquidity flows and fragility of business enterprises. *Journal of Monetary Economics*, 50(6):1215–1241.

Eisfeldt, A. and Rampini, A. (2006). Capital reallocation and liquidity. *Journal of Monetary Economics*, 53(3):369–399.

Fujita, S. and Ramey, G. (2009). The cyclicality of separation and job finding rates. *International Economic Review*, 50(2):415–430.

Greenstone, M., Mas, A., and Nguyen, H.-L. (2020). Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and 'normal' economic times. *American Economic Journal: Economic Policy*, 12(1):200–225.

Gürkaynak, R., Sack, B., and Swanson, E. (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*.

Herrera, A. M., Kolar, M., and Minetti, R. (2011). Credit reallocation. *Journal of Monetary Economics*, 58(6-8):551–563.

Hubbard, G., Kuttner, K., and Palia, D. (2002). Are there bank effects in borrowers' costs of funds? evidence from a matched sample of borrowers and banks. *The Journal of Business*, 75(4):559–581.

Ivashina, V. and Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics*, 97(3):319–338.

Jarociński, M. and Karadi, P. (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Jiménez, G., Ongena, S., Peydró, J.-L., and Saurina, J. (2014). Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? *Econometrica*, 82(2):463–505.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182.

Kerssenfischer, M. (2019). Information effects of euro area monetary policy: New evidence from high-frequency futures data. Technical report, Deutsche Bundesbank Discussion Paper.

Khwaja, A. I. and Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, 98(4):1413–42.

Kiyotaki, N. and Moore, J. (1997). Credit cycles. *Journal of Political Economy*, 105(2):211–248.

Lilien, D. and Hall, R. (1986). Cyclical fluctuations in the labor market. *Handbook of Labor Economics*, 2(Part C):1001–1035.

Mazet-Sonilhac, C. (2020). Information friction in credit markets. Technical report, Working Paper.

Nakashima, K. and Takahashi, K. (2018). The real effects of bank-driven termination of relationships: Evidence from loan-level matched data. *Journal of Financial Stability*, 39:46–65.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Rogerson, R. and Shimer, R. (2011). Search in macroeconomic models of the labor market. *Handbook of Labor Economics*, 4:619–700.

A. Tables and Figures

Table 3.1: Summary Statistics: Aggregate Results

This table reports summary statistics aggregated at the quarter level for the period 1999Q1-2016Q4. All credit variables are in thousands of Euro unless specified otherwise and are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. We display the mean and standard deviation for all variables over the full sample period in addition to the first (1999) and last (2016) complete years in our sample.

	Full sample		1999		2016	
	Mean	SD	Mean	SD	Mean	SD
Number of banks	447.02	60.85	541.98	3.89	360.49	6.57
Number of firms	256271.62	44836.66	182126.04	4433.08	301378.04	1428.21
Number of credit relationships	345678.89	50645.87	264776.49	5594.82	399220.31	2253.39
Aggregate credit exposure (Eur Bn)	358.13	60.52	258.77	9.38	408.86	5.20
Number of relationships per firm	1.36	0.05	1.45	0.00	1.32	0.00
Number of relationships per bank	802.50	219.30	488.57	14.52	1107.74	25.90
Fraction of firms with 1 bank	80.30	1.95	76.41	0.15	81.47	0.05
Fraction of firms with 2 banks	12.19	0.78	13.70	0.03	11.81	0.03
Fraction of long-term only	45.71	1.26	42.42	0.61	44.91	0.51
Fraction of short-term only	30.83	3.89	37.77	0.29	25.96	0.34
Fraction of short-term + long-term only	14.20	1.60	16.75	0.25	13.71	0.19
Fraction of undrawn only	9.26	4.87	3.06	0.31	15.42	0.25
Relationship duration (in quarters)	14.64	4.79	4.91	0.79	21.52	0.32
Credit exposure per match	1032.93	48.84	977.08	16.14	1024.11	7.29
Short-term debt per match	214.05	71.33	329.22	4.02	149.24	0.46
Long-term debt per match	413.96	51.85	334.44	7.29	454.65	3.23
Undrawn credit line per match	396.14	55.37	311.60	7.88	406.95	6.01
Share of long-term credit per match	50.05	1.22	48.57	0.57	49.73	0.28
Share of drawn credit per match	81.45	6.10	89.45	0.32	74.99	0.20
Fraction of credit to new entrants	4.15	1.10	5.27	0.21	3.14	0.34
Fraction of credit to incumbents	95.85	1.15	94.86	0.50	96.82	0.54
Average credit per entering firm / incumbent	57.48	8.26	58.67	2.29	50.27	1.55
Average credit per exiting firm / incumbent	44.77	6.68	50.00	4.52	38.28	1.38
Creation flow	23407.40	1950.81	22502.74	877.01	23812.29	828.83
Destruction flow	21497.35	1992.70	17672.60	312.80	23223.42	417.51
Net flow	1910.05	2288.60	4830.14	568.97	588.86	1124.07
Excess reallocation	42464.36	3618.07	35345.20	625.59	46129.19	528.46
Creation rate	6.94	1.04	8.63	0.30	5.98	0.22
Destruction rate	6.32	0.52	6.78	0.14	5.83	0.10
Net flow rate	0.62	0.75	1.85	0.20	0.15	0.28
Excess reallocation rate	12.51	1.14	13.56	0.27	11.58	0.10
Fraction of switching firms	0.42	0.08	0.54	0.01	0.35	0.01
Firm entry rate	4.57	0.54	4.99	0.21	4.14	0.09
Firm exit rate	3.79	0.21	3.47	0.03	3.84	0.02
Firm entry / firm creation	22.82	2.39	27.25	1.68	19.70	0.44

Table 3.2: Cyclical Properties of Credit Relationship Flows

This table reports the results for auto-correlation, standard deviation of detrended credit relationships flows and their correlation with respect to the log-growth of GDP, total credit, and relationship capital, over the period 1999-2016. The top panel shows results for flows, in levels, while the bottom panel shows the results in rates. All flow variables are detrended using an HP filter with a smoothing parameter of 1600. All nominal credit variables are deflated using the French GDP deflator, deseasonalized using the X-13 seasonal adjustment procedure, and smoothed based on MA(-1, 1).

	Autocor(x)	Stdev(x)	cor(x, GDP)	cor(x, Total credit)	cor(x, Relationship capital)
			A. Levels		
Creation flows	0.749	0.044	0.354	0.445	0.629
Destruction flows	0.673	0.026	-0.374	-0.155	-0.278
Net flows	0.754	0.051	0.494	0.458	0.678
Reallocation	0.701	0.050	0.107	0.300	0.394
Excess reallocation	0.683	0.025	-0.278	-0.057	-0.045
			B. Rates		
Creation flows	0.730	0.003	0.432	0.474	0.639
Destruction flows	0.589	0.001	-0.261	-0.138	-0.258
Net flows	0.738	0.004	0.498	0.485	0.683
Reallocation	0.677	0.004	0.288	0.378	0.479
Excess reallocation	0.604	0.003	-0.141	-0.013	0.008

Table 3.3: Cyclical Properties of Credit Relationship Flows: Cross-sectional Decomposition
This table reports the results for auto-correlation; standard deviation of detrended credit relationships flows decomposed by (i) credit size, (ii) credit type, and (iii) relationship duration; and their cross-correlation with log-growth GDP, total credit, and relationship capital, over the period 1999-2016. All flow variables are detrended using an HP filter with a smoothing parameter of 1600. All nominal credit variables are deflated using the French GDP deflator, deseasonalized using the X-13 seasonal adjustment procedure, and smoothed based on MA(-1, 1).

	Autocor(x)	Stdev(x)	cor(x, GDP)	cor(x, Total credit)	cor(x, Rel. capital)
			A. Credit size		
Creation flows: Small	0.841	0.003	0.493	0.478	0.709
Destruction flows: Small	0.674	0.002	-0.296	-0.088	-0.141
Net flows: Small	0.786	0.004	0.480	0.546	0.645
Creation flows: Medium	0.863	0.003	0.501	0.485	0.644
Destruction flows: Medium	0.745	0.001	-0.332	-0.106	0.123
Net flows: Medium	0.852	0.003	0.622	0.511	0.569
Creation flows: Large	0.886	0.003	0.562	0.508	0.590
Destruction flows: Large	0.660	0.001	0.040	-0.059	0.271
Net flows: Large	0.858	0.003	0.561	0.542	0.506
Creation flows: Very large	0.906	0.003	0.497	0.614	0.562
Destruction flows: Very large	0.580	0.001	0.153	0.283	0.200
Net flows: Very large	0.818	0.002	0.476	0.534	0.524
			B. Credit type		
Creation flows: Long-term	0.879	0.003	0.656	0.578	0.610
Destruction flows: Long-term	0.659	0.001	0.315	0.282	0.098
Net flows: Long-term	0.826	0.003	0.592	0.519	0.650
Creation flows: Short-term	0.841	0.004	0.178	0.225	0.412
Destruction flows: Short-term	0.838	0.004	-0.516	-0.386	-0.086
Net flows: Short-term	0.826	0.005	0.523	0.465	0.396
Creation flows: Credit line	0.627	0.014	0.356	0.120	0.367
Destruction flows: Credit line	0.600	0.008	-0.430	0.102	-0.189
Net flows: Credit line	0.616	0.016	0.497	0.051	0.393
			C. Relationship duration		
Destruction flows: < 1 year	0.782	0.003	-0.370	-0.269	-0.437
Destruction flows: $1 < 2$ years	0.642	0.002	0.178	-0.064	-0.058
Destruction flows: $2 < 5$ years	0.755	0.002	0.183	0.189	0.111
Destruction flows: ≥ 5 years	0.712	0.002	-0.252	0.247	0.165

Table 3.4: Cyclical Properties of Aggregate Variables

This table reports the results for auto-correlation; standard deviation; and correlation of GDP, total credit, and relationship capital, over the period 1999-2016. In the top panel of the table, results are based on the log-growth of the variables. In the bottom panel, results are based on log-deviations from HP trends, obtained using an HP filter with a smoothing parameter of 1600. All nominal credit variables are deflated using the French GDP deflator, deseasonalized using the X-13 seasonal adjustment procedure, and smoothed based on MA(-1, 1).

	A. Log-growth				
	Autocor(x)	Stdev(x)	cor(x, GDP)	cor(x, Total credit)	cor(x, Relationship capital)
GDP	0.868	0.004	1.000	0.378	0.443
Total credit	0.831	0.013	0.378	1.000	0.640
Relationship capital	0.769	0.006	0.443	0.640	1.000
Average credit	0.719	0.010	0.230	0.904	0.250
	B. Cyclical deviations				
	Autocor(x)	Stdev(x)	cor(x, GDP)	cor(x, Total credit)	cor(x, Relationship capital)
GDP	.926	0.009	1.000	0.480	0.558
Total credit	0.935	0.029	0.480	1.000	0.707
Relationship capital	0.908	0.010	0.558	0.707	1.000
Average credit	0.923	0.023	0.364	0.954	0.462

Table 3.5: Variance Decomposition: Intensive vs. Extensive Margins

This table reports the results of variance decompositions of aggregate credit fluctuation over the period 1999-2016. The intensive/extensive margin decompositions are derived based on first-differences (Panel A) and log-deviations from trend (Panel B) obtained from an HP filter with a smoothing parameter of 1600. All nominal credit variables are deflated using the French GDP deflator, deseasonalized using the X-13 seasonal adjustment procedure, and smoothed based on MA(-1, 1).

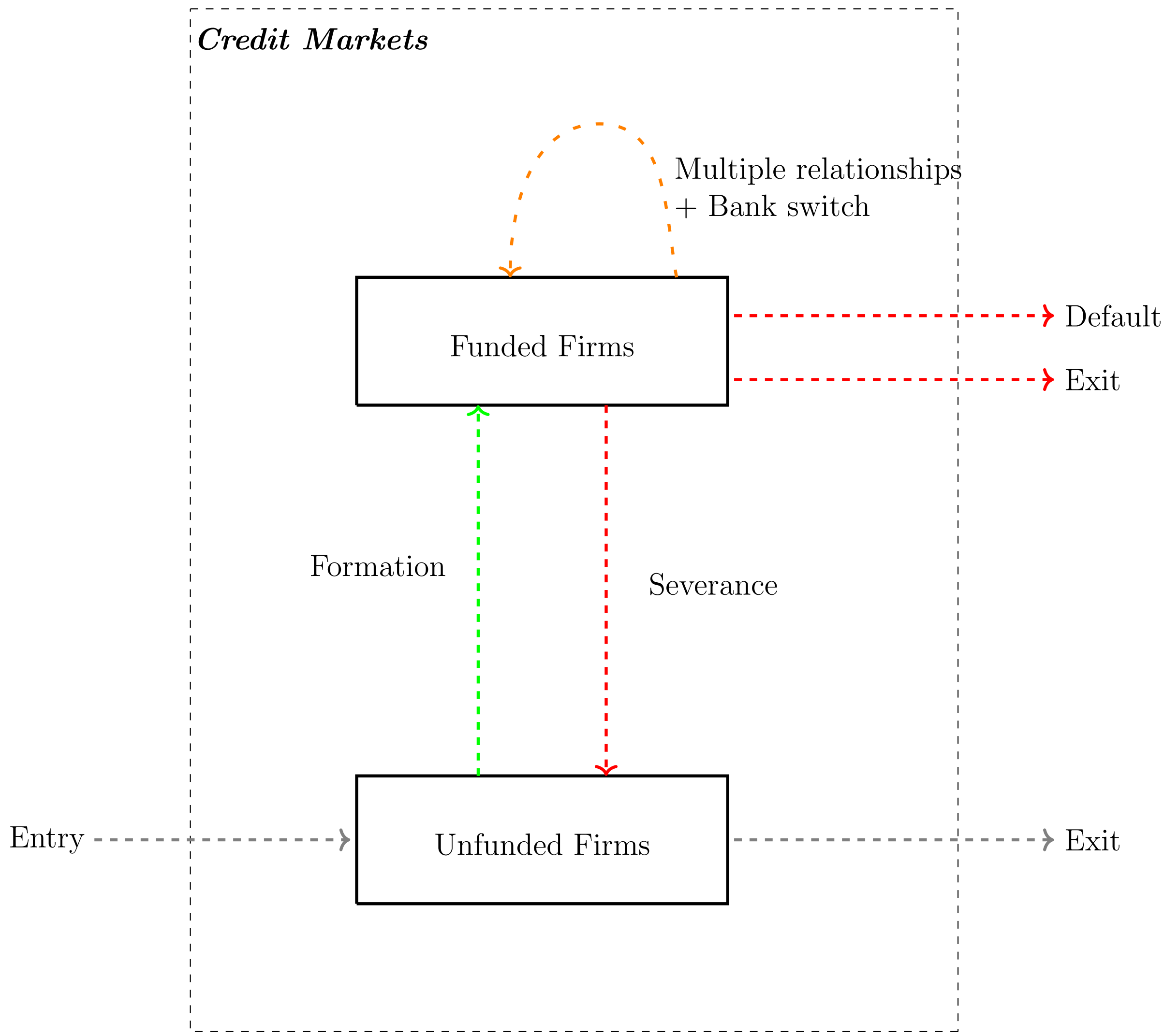
A. Decomposition 1

First-Difference Approach	Intensive Margin	Extensive Margin	
	0.73	0.27	
		Creation Flows	Destruction Flows
		0.43	-0.16
HP Filter Approach	Intensive Margin	Extensive Margin	
	0.76	0.22	
		Creation Flows	Destruction Flows
		0.23	-0.03

B. Decomposition 2

First-Difference Approach	Intensive Margin	Extensive Margin	
	0.54	0.46	
	Incumbent effect	New bank-firm effect	Severed bank-firm effect
	0.54	0.62	-0.17
HP Filter Approach	Intensive Margin	Extensive Margin	
	0.46	0.40	
	Incumbent effect	New bank-firm effect	Severed bank-firm effect
	0.46	0.57	-0.17

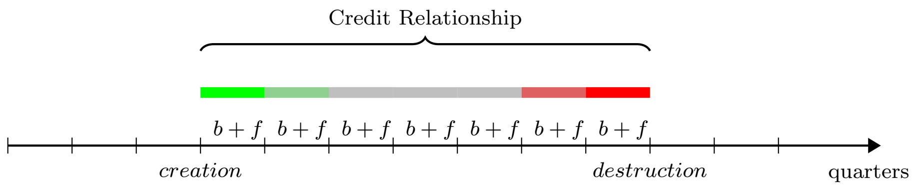
Fig. 3.1. The Flow Approach to Credit Markets



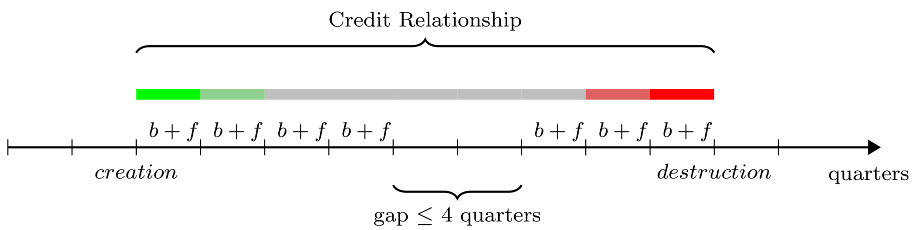
Note: This figure displays the multiple forms of flows associated with unfunded and funded firms within credit markets.

Fig. 3.2. Credit Relationships: Concepts and Measurements

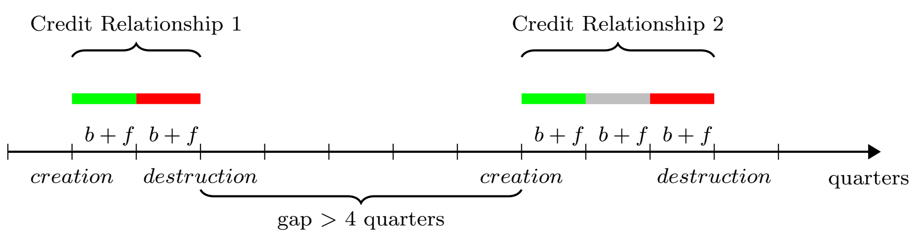
(a) Case 1: **Contiguous credit relationship (no gap)**



(b) Case 2: **Contiguous credit relationship (gap below 4 quarters)**



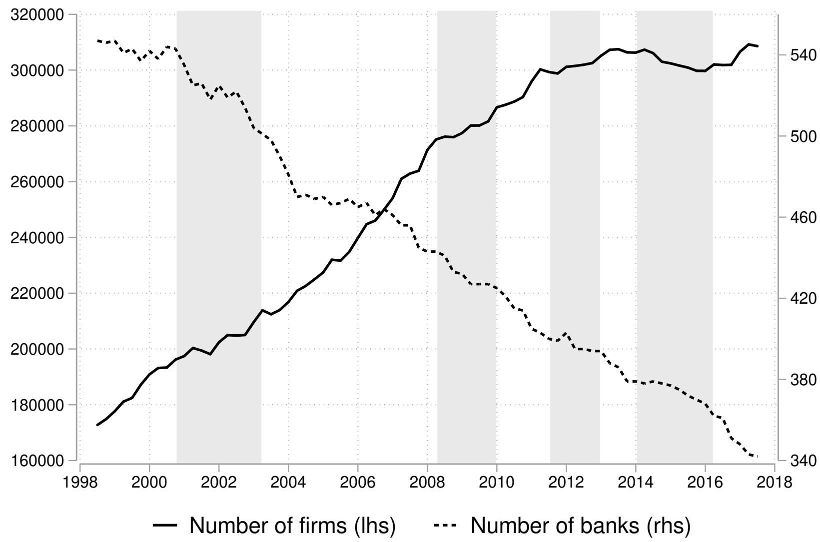
(c) Case 3: **Non-contiguous credit relationships (gap exceeding 4 quarters)**



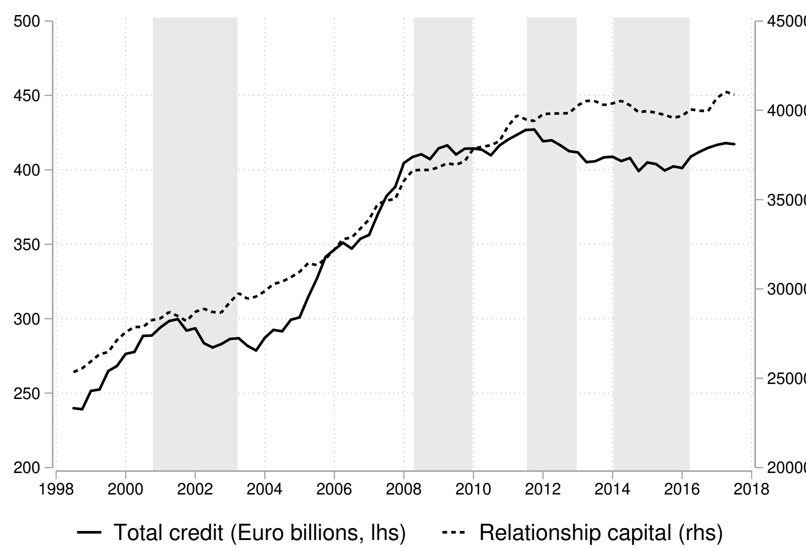
Notes: These figures represent potential situations for bank-firm (b+f) match data entries and the corresponding definitions for credit relationships and gross flows. We consider that a credit relationship is "contiguous" as long as the data entries are available with a reporting gap below 4 quarters (cases (a) and (b)). When the reporting gap is above 4 quarters, we consider that the bank-firm entries generate 2 non-contiguous credit relationships with independent creation and destruction dates (case (c)).

Fig. 3.3. Evolution of Banks, Firms, Bank-firm Relationships, and Credit

(a) **Total banks and firms**



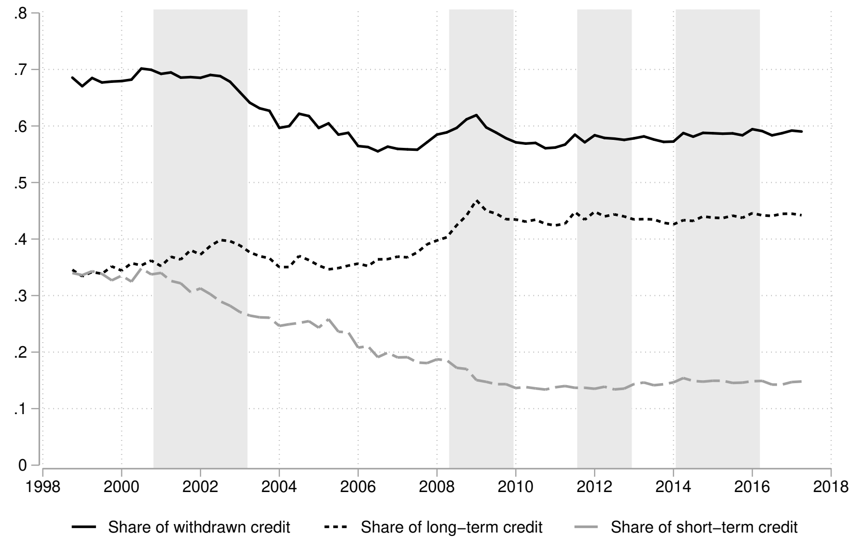
(b) **Aggregate credit and number of bank-firm relationships**



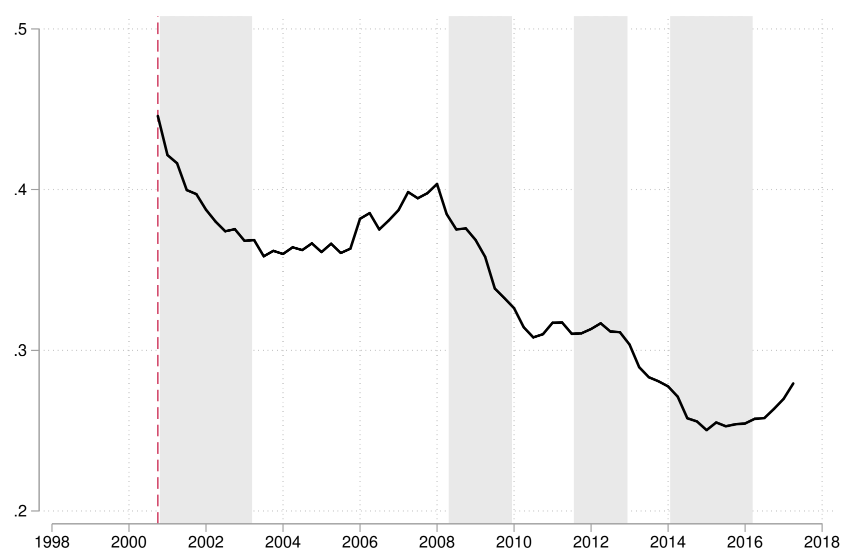
Notes: Panel (a) shows the evolution of the number of unique firms and active banks. Only those banks and firms involved in credit relationships with credit exposure above the reporting threshold are taken into account. Panel (b) shows the evolution of aggregate bank credit (solid line) and the number of bank-firm relationships (dashed line), with credit exposure above the reporting threshold. The sample period is 1999-2017 . All nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Gray-shaded areas correspond to recession periods.

Fig. 3.4. Share of Credit Relationships by Type and Duration

(a) **Decomposition by type/maturity**



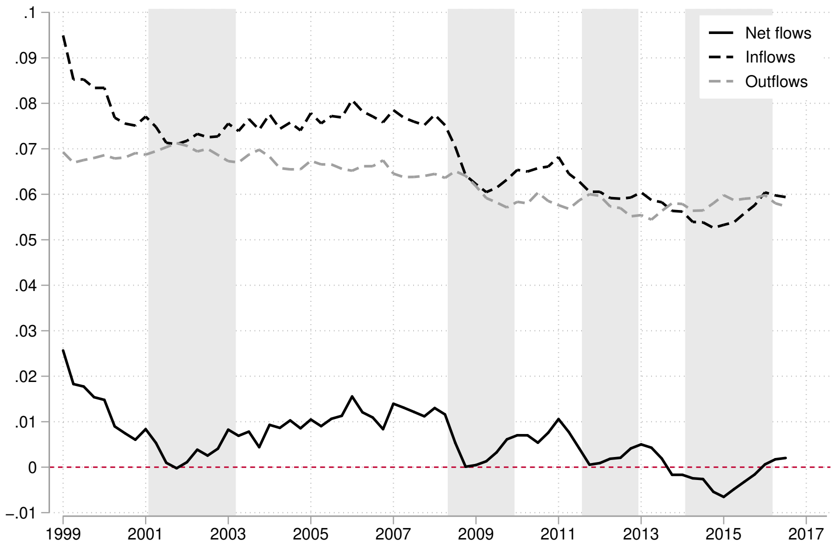
(b) **Share of relationships below two-year duration**



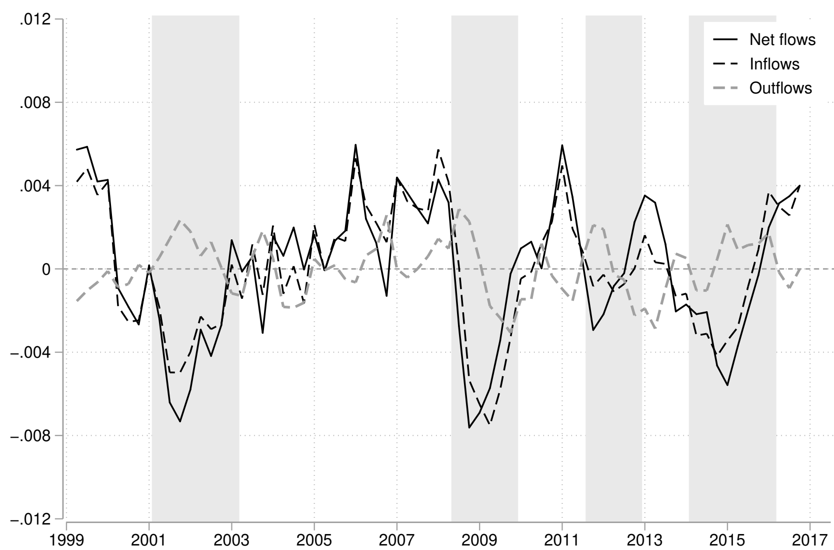
Notes: Panel (a) shows the share of credit relationship per type and maturity over the sample period. Panel (b) shows the share of credit relationship by duration. Results are based on relationships above the reporting threshold (adjusted for inflation) and reported over the period 1999-2017. Gray-shaded areas correspond to recession periods.

Fig. 3.5. Credit Relationship Flows

(a) **Raw net and gross flows**



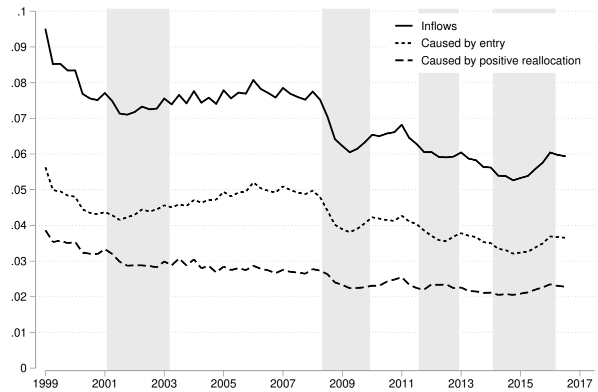
(b) **Cyclical deviations of net and gross flows**



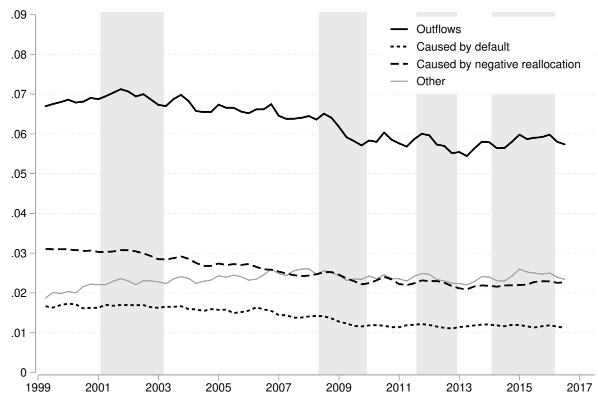
Notes: Panel (a) shows raw net (solid black line) and gross flows of credit relationships. Gross creation flows (inflows) are reported in dashed black line, while gross destruction flows (outflows) are reported in dashed gray line. Panel (b) shows the time series for cyclical deviations corresponding to the same three variables after applying an HP filter with a smoothing parameter of 1600. Our sample period is 1999-2016. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Gray-shaded areas correspond to recession periods.

Fig. 3.6. Sources of Relationship Creation and Destruction

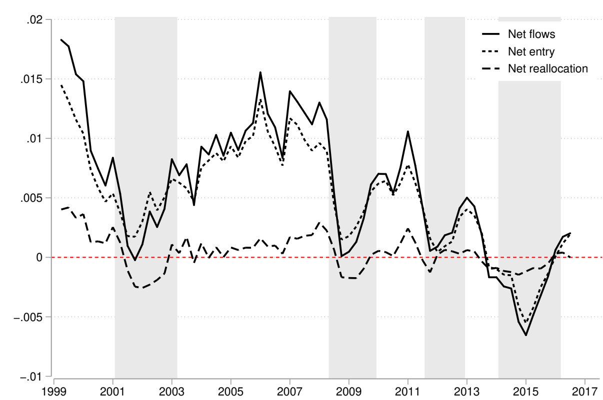
(a) **Decomposition of inflows**



(b) **Decomposition of outflows**

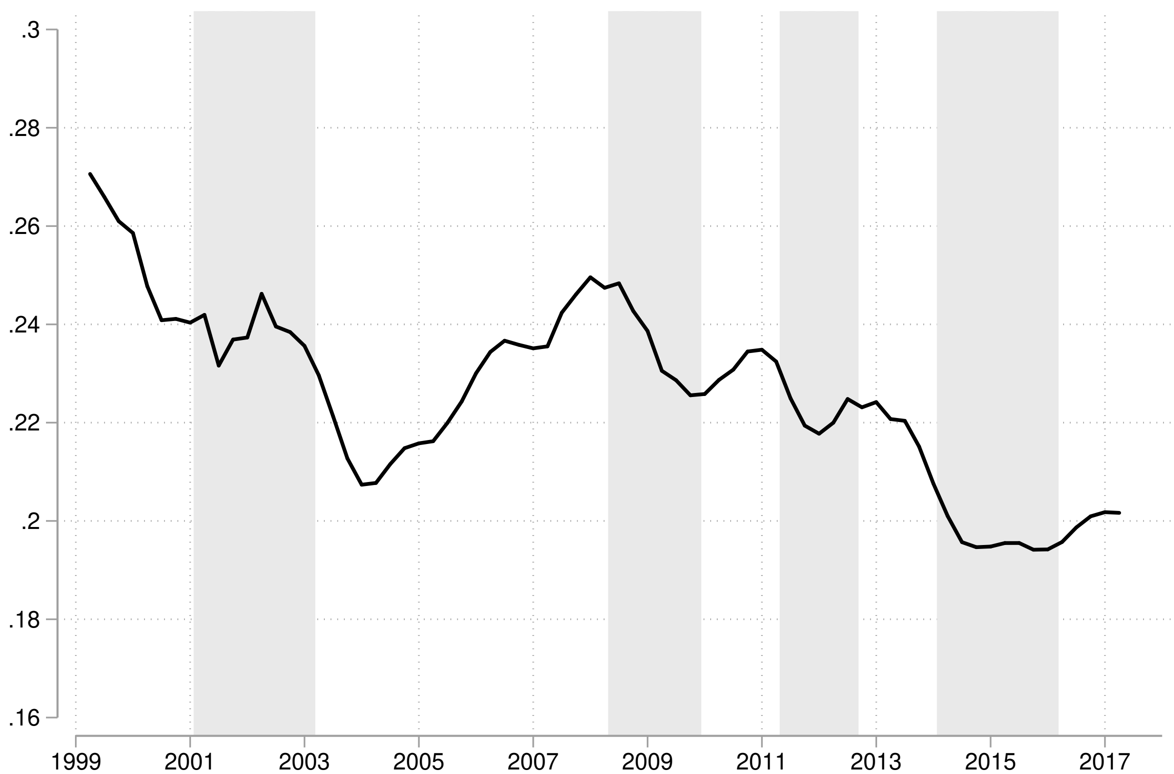


(c) **Decomposition of net flows**



Notes: This figure shows the decomposition of raw creation (Panel (a)), destruction (Panel (b)), and net (Panel (c)) flows due to firms (i) entering or exiting the relationship, (ii) switching borrowers, or (iii) experiencing multi-bank relationship gains or losses. Our sample period is 1999-2016. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Gray-shaded areas correspond to recession periods.

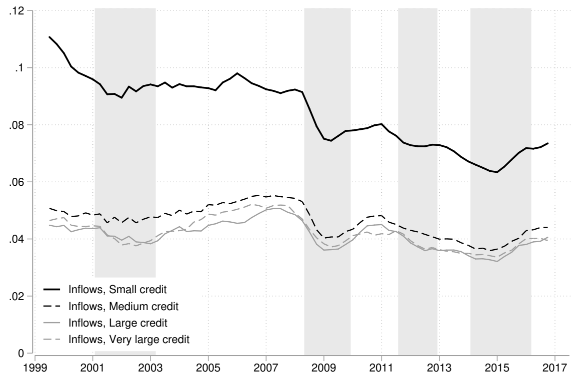
Fig. 3.7. First Credit Relationship and Firm Entry



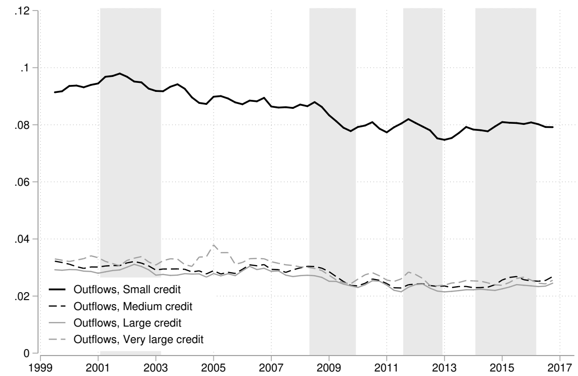
Notes: This figure reports the ratio of first-time borrowers over total number of newly created firms. Our sample period is 1999-2016. Results are based on relationships above the reporting threshold (adjusted for inflation). Gray-shaded areas correspond to recession periods.

Fig. 3.8. Gross Flows, by Credit Size, Type, and Duration

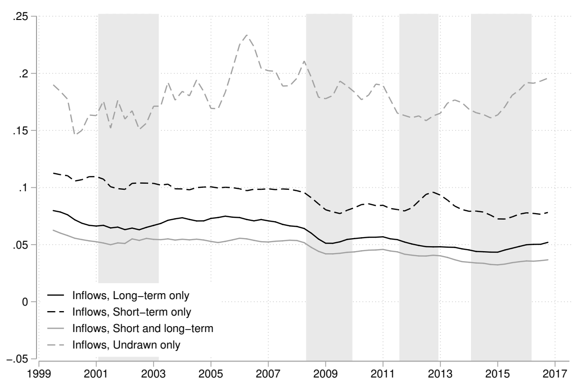
(a) **Inflows by credit size**



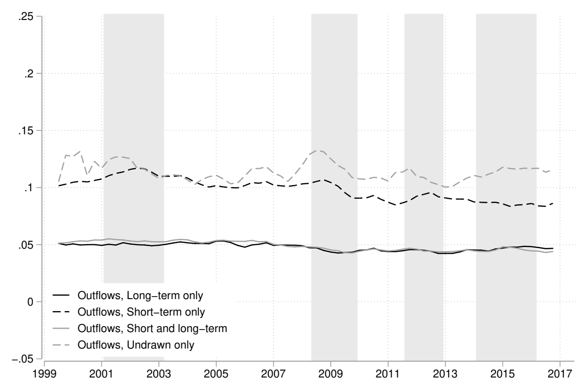
(b) **Outflows by credit size**



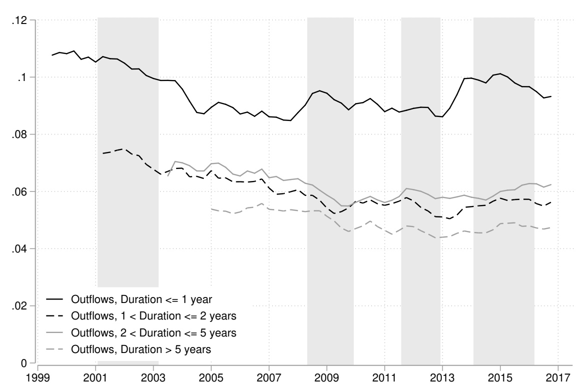
(c) **Inflows by credit type**



(d) **Outflows by credit type**

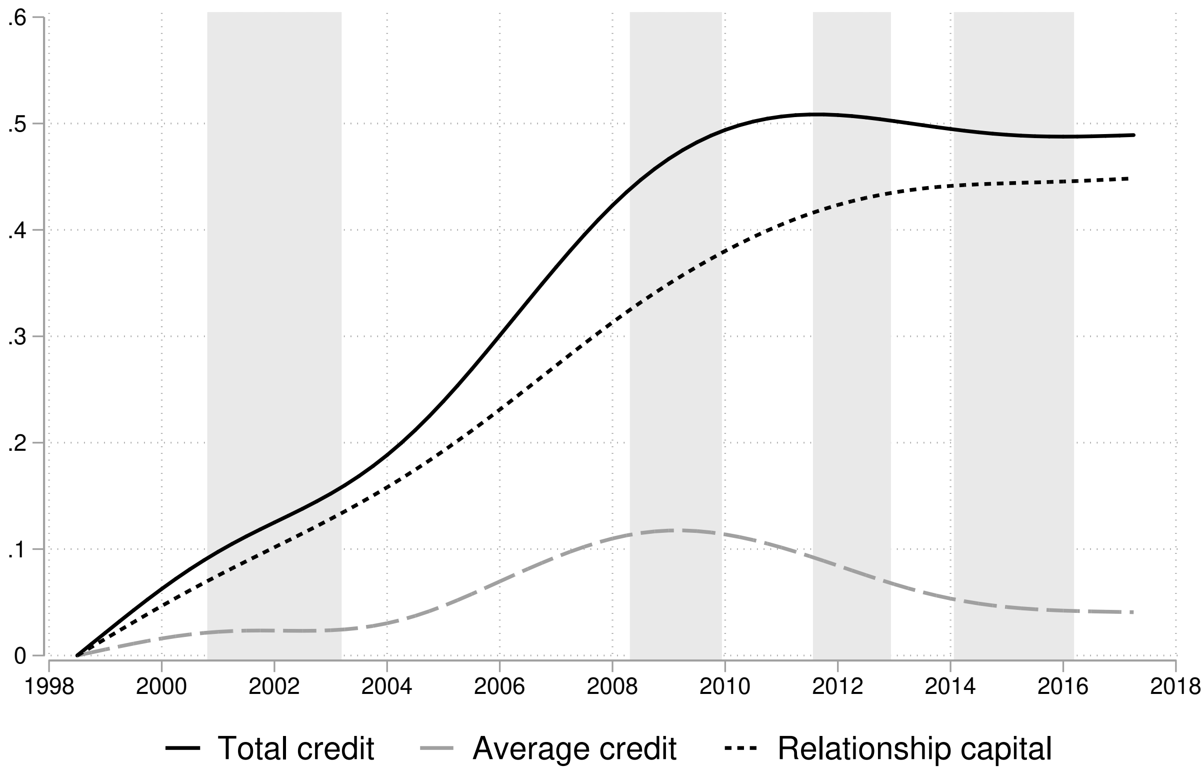


(e) **Outflows by relationship duration**



Notes: This figure shows the decomposition of raw creation (left panels) and destruction (right panels), by credit size (Panels (a) & (b)), by type (Panels (c) & (d)), and by relationship duration (Panel (e), for outflows only). Our sample period is 1999-2016. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Gray-shaded areas correspond to recession periods.

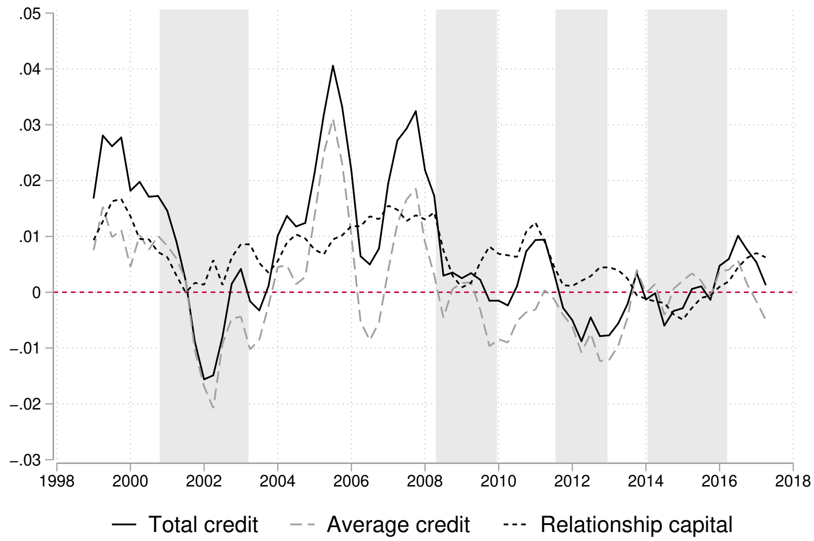
Fig. 3.9. Extensive vs. Intensive Margins: Long-run Trends



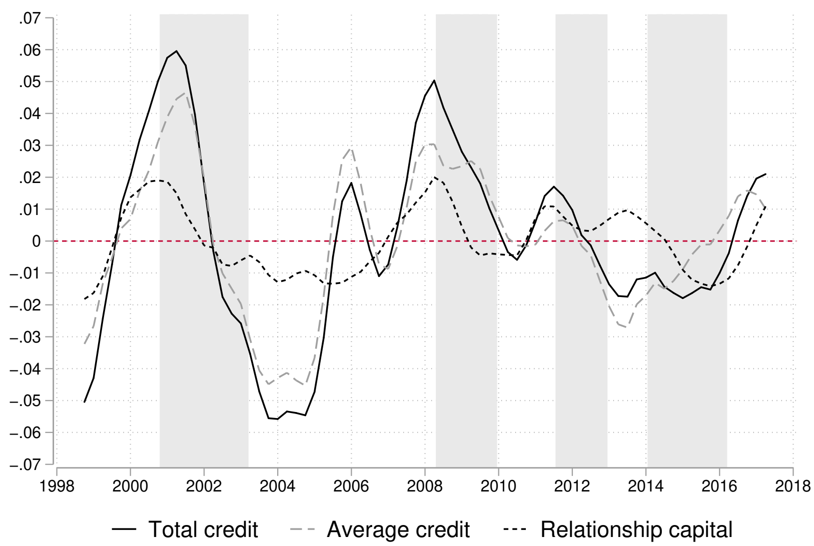
Notes: This figure reports the trends associated with aggregate credit, average credit, and relationship capital, obtained through the simple decomposition 1. The trends are extracted using an HP filter with a smoothing parameter of 1600. Nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Our sample period is 1999-2017. Gray-shaded areas correspond to recession periods.

Fig. 3.10. Extensive vs. Intensive Margins of Credit – Decomposition 1

(a) **First-difference approach**



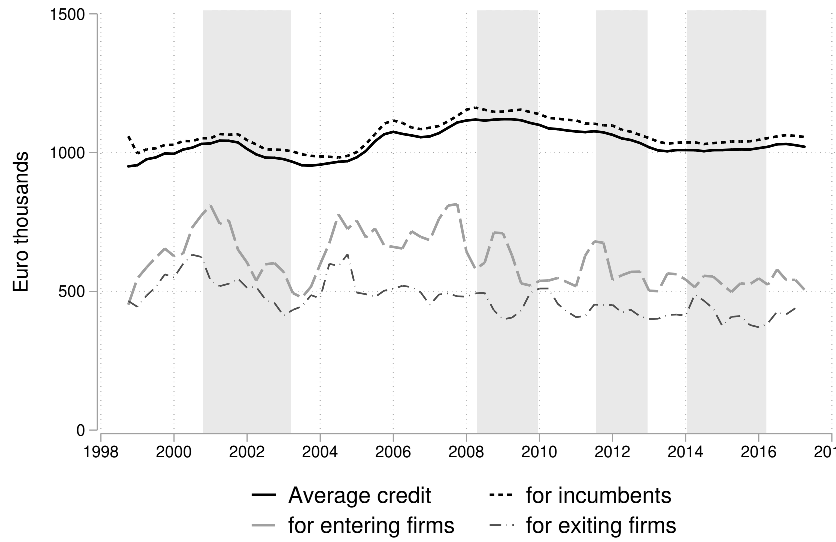
(b) **HP Filter approach**



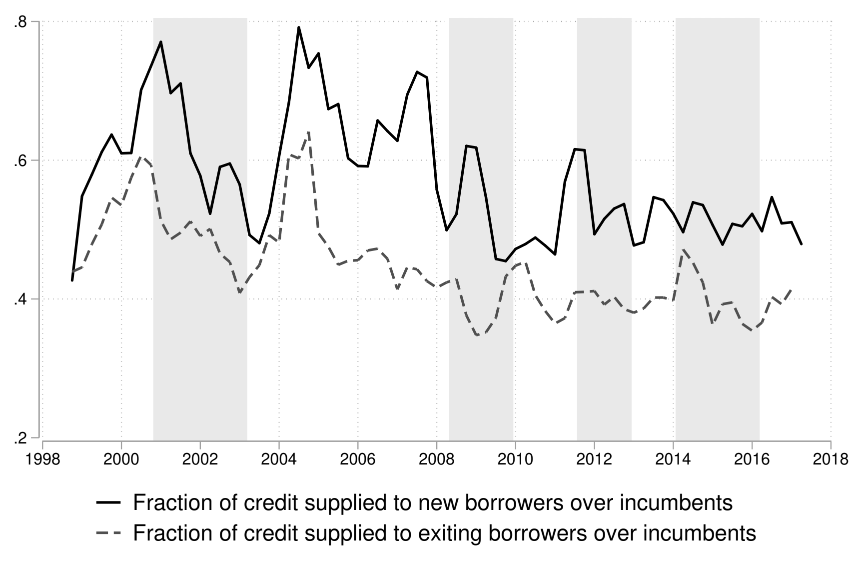
Notes: These figures show the log-growth dynamics (Panel (a)) and cyclical deviations (in log, Panel (b)) of aggregate credit (black solid line), average credit per relationship (gray dashed line), and relationship capital (black dashed line), obtained through the simple decomposition 1. Nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Our sample period is 1999-2016. Gray-shaded areas correspond to recession periods.

Fig. 3.11. Credit for Incumbents vs. Entering and Exiting Firms

(a) **Average credit for incumbents vs. entering and exiting firms**



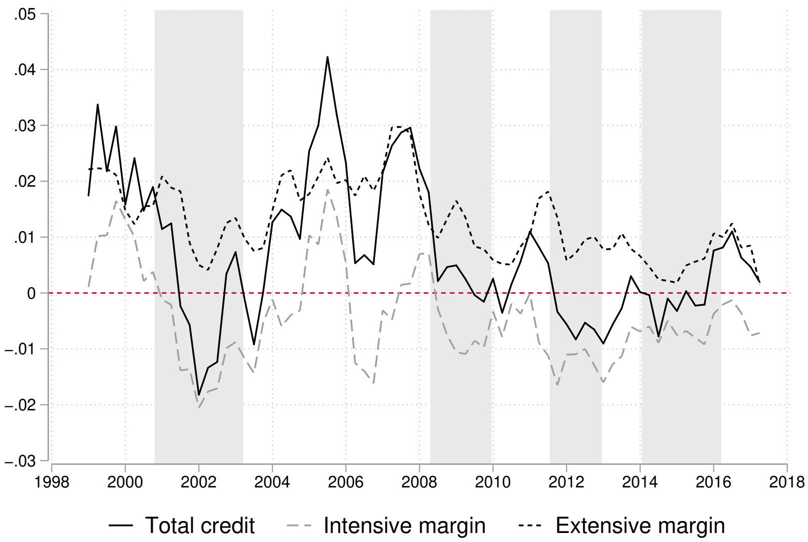
(b) **Credit supplied to entering and exiting firms, as a fraction of incumbent credit**



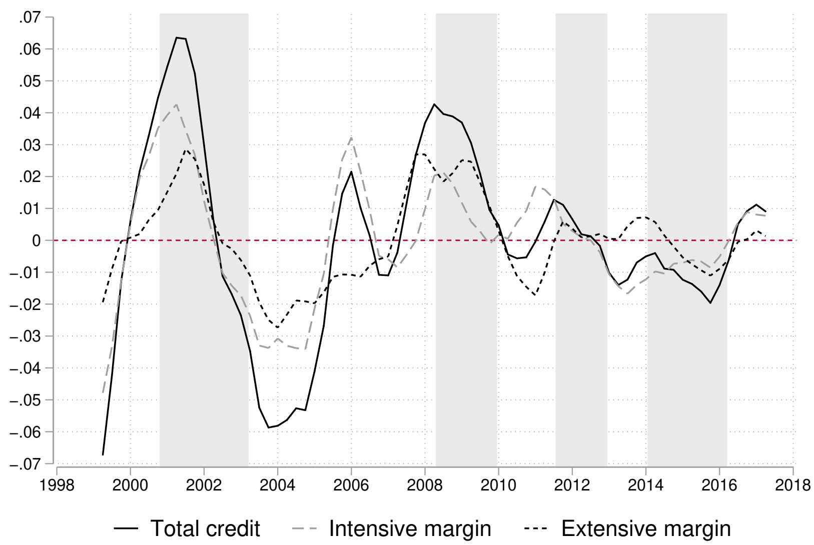
Notes: Panel (a) shows the time series of aggregate average credit per relationship (solid black line), in addition to the average credit supplied to (i) incumbent borrowers (black dotted line), (ii) new borrowers (light gray dashed line), and (iii) exiting borrowers (dark gray dashed line). Panel (b) shows the time series of the ratio of (i) average credit supplied to new borrowers over average credit supplied to incumbents (solid line) and (ii) average credit (previously) supplied to exiting borrowers over average credit supplied to incumbents (dashed line). Our sample period is 1999-2016. Results are based on relationships above the reporting threshold (adjusted for inflation). All nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator from the FRED database. Gray-shaded areas correspond to recession periods.

Fig. 3.12. Extensive vs. Intensive Margins of Credit – Decomposition 2

(a) **First-difference Approach**



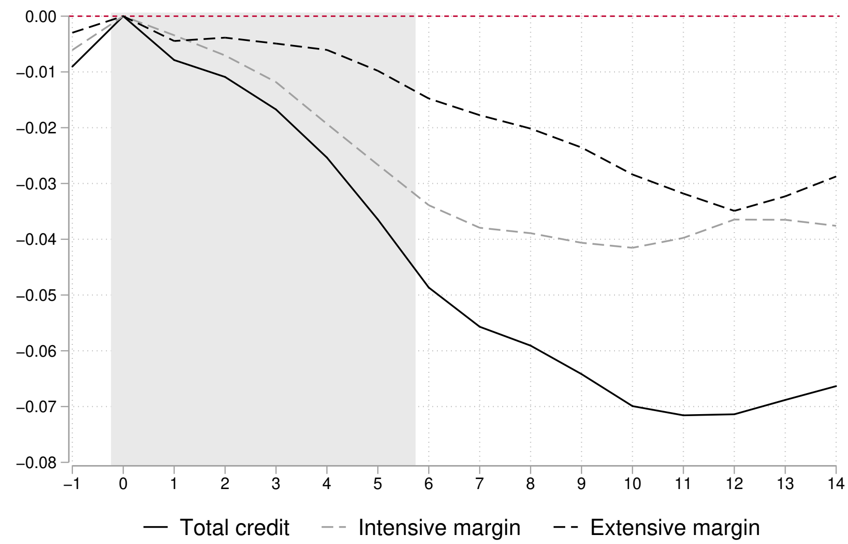
(b) **HP Filter Approach**



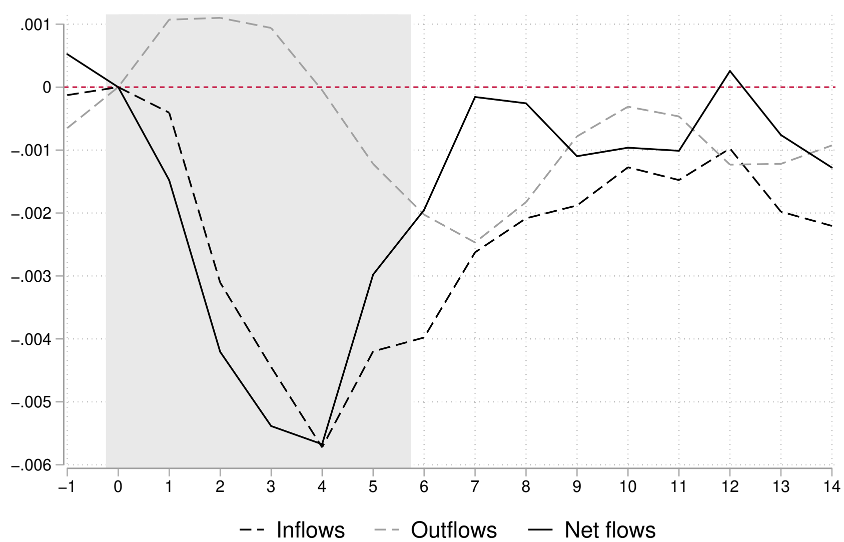
Notes: These figures show the time series of the log-growth dynamics (Panel (a)) cyclical deviations (in log, Panel (b)) for total credit (solid black line), intensive margin (gray dashed line), and extensive margin (black dashed line), obtained through the refined decomposition 2. The intensive margin is the change in the average credit supplied to incumbents multiplied by the number of incumbents. The extensive margin is the number of new relationships multiplied by the average credit supplied to new firms minus the number of exiting relationships multiplied by the average credit supplied to exiting firms. Nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Our sample period is 1999-2016. Gray-shaded areas correspond to recession periods.

Fig. 3.13. Anatomy of a Crisis: Unconditional Patterns

(a) **Aggregate variables: credit vs. intensive and extensive margins**

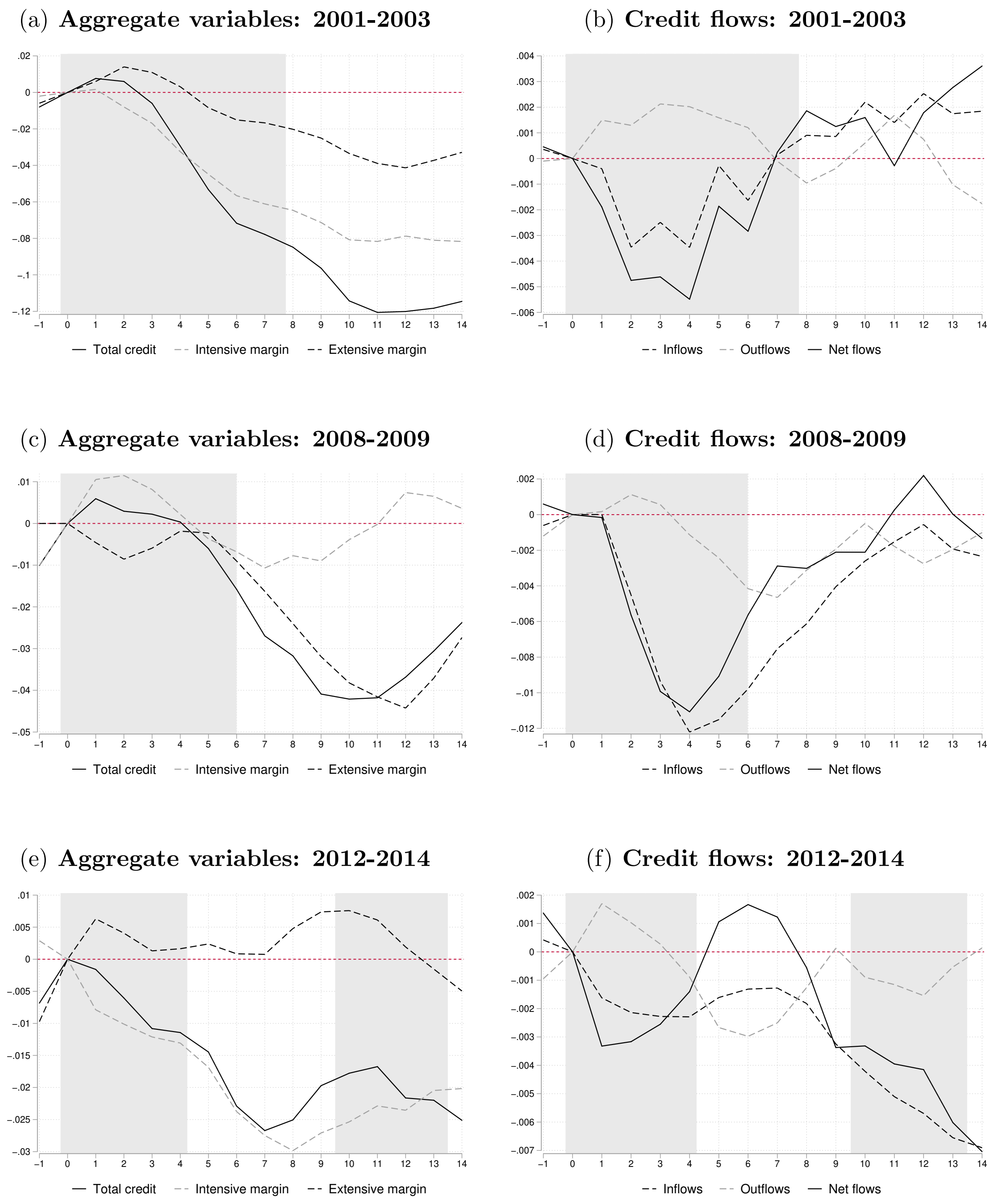


(b) **Credit relationship flows: net vs. gross**



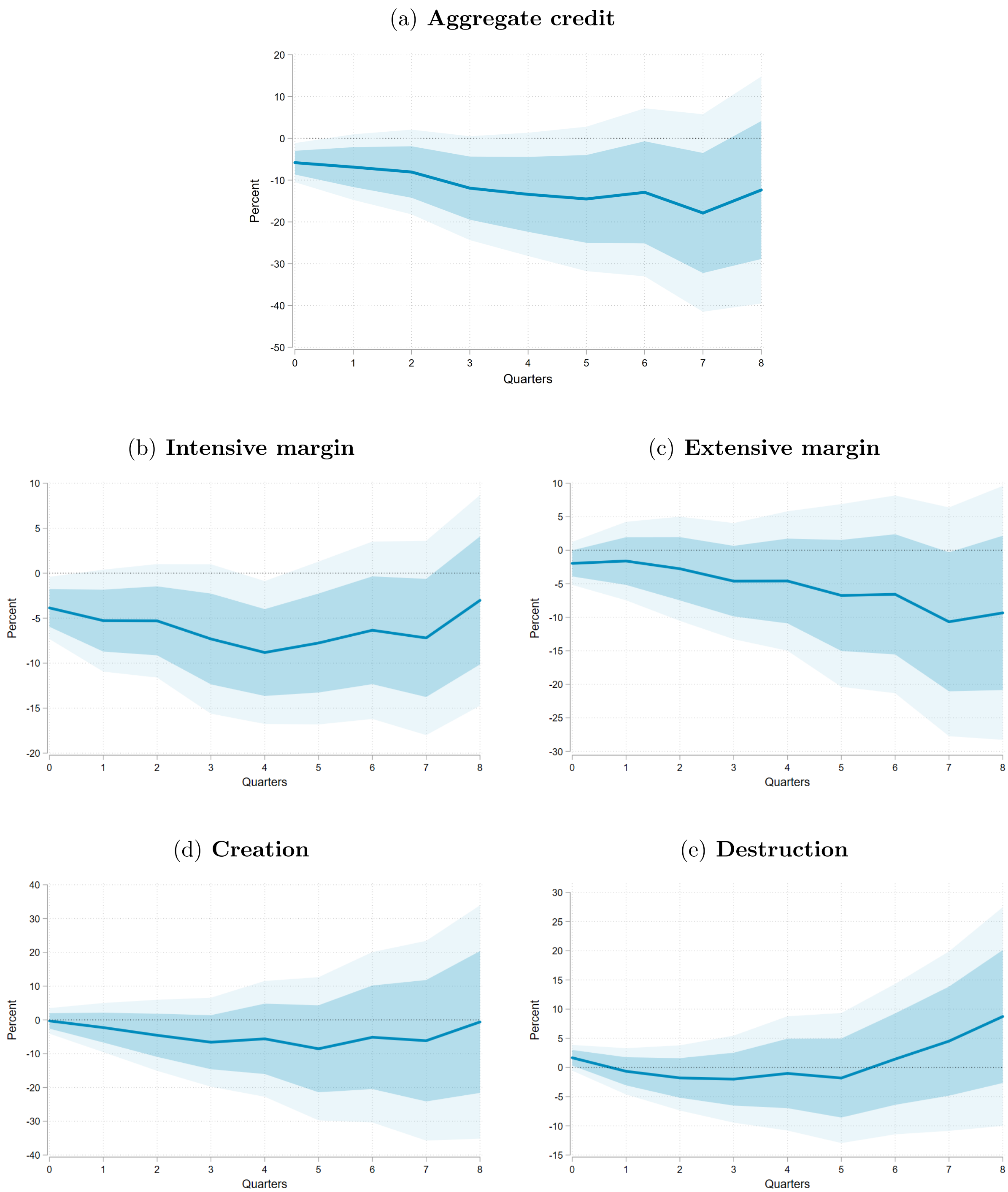
Notes: These figures report the unconditional evolution of (i) aggregate credit, intensive, and extensive margins capital (top panel), and that of (ii) net and gross flows (bottom panel) over the fourteen quarters following the onset of a recession. The intensive/extensive margins are constructed based on the refined credit decomposition 2 specified in equation (3.10). The aggregate credit dynamics reported are based on the sum of the extensive and intensive margins. All variables are normalized to 0 based on the timing of the pre-recession peak for aggregate credit, and reported in terms of log-deviations from their corresponding HP trend obtained with a smoothing parameter of 1600. Gray-shaded areas correspond to the average recession period.

Fig. 3.14. Anatomy of a Crisis: Details



Notes: These figures report the evolution of (i) aggregate credit, intensive, and extensive margins (left-hand side panels), and (ii) net and gross flows (right-hand side panels) over the fourteen quarters following the onset of each recession. The intensive/extensive margins are constructed based on the refined credit decomposition 2 specified in equation (3.10). The aggregate credit dynamics reported are based on the sum of the intensive and extensive margins. Due to their proximity, the recessions of 2012-2013 and 2014-2016 are shown combined in Panels (e) and (f). All variables are normalized to 0 based on the timing of the pre-recession peak for aggregate credit, and reported in terms of log-deviations from their corresponding HP trend obtained with a smoothing parameter of 1600. Gray-shaded areas correspond to recession periods.

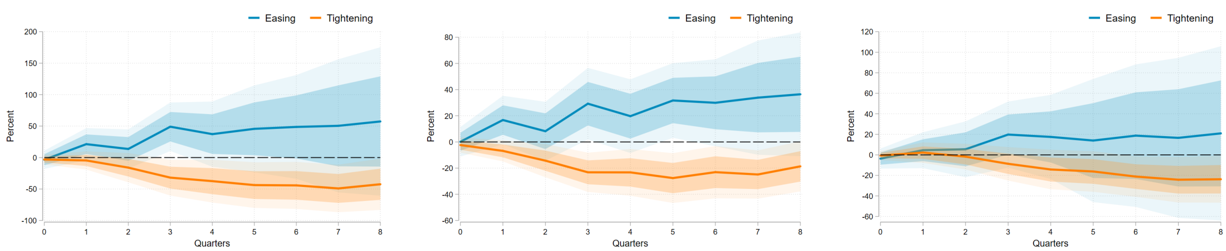
Fig. 3.15. Monetary Policy Transmission and Credit



Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, (c) extensive margin, and the corresponding (d) creation and (e) destruction components. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.15) and the "purified" monetary policy surprises from Jarociński and Karadi (2020). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The blue-shaded areas correspond to the 68% (dark blue) and 90% (light blue) confidence intervals constructed using Newey and West (1987) standard errors.

Fig. 3.16. Monetary Policy Transmission and Credit – Easing vs. Tightening

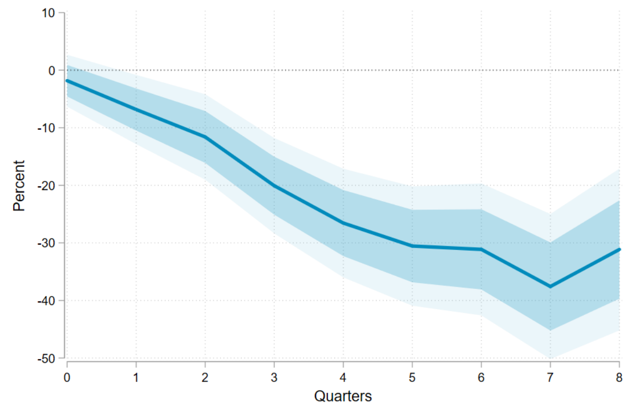
(a) **Aggregate credit** (b) **Intensive margin** (c) **Extensive margin**



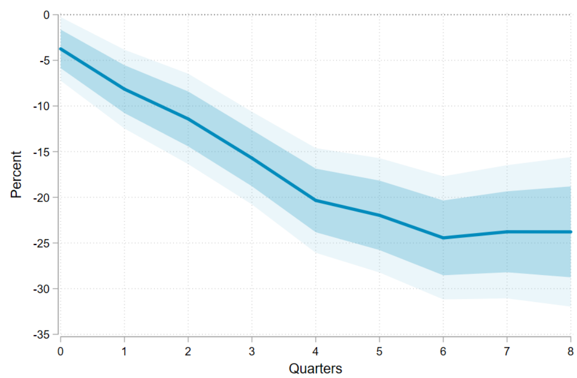
Notes: These figures illustrate impulse responses to a one percentage point contractionary (orange) and expansionary (blue) monetary policy shock for (a) aggregate credit, (b) intensive margin, and (c) extensive margin. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.15) and the "purified" monetary policy surprises from Jarociński and Karadi (2020). The local projections are estimated following the specification described in equation (3.16). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The shaded areas correspond to the 68% (dark color) and 90% (light color) confidence intervals constructed using Newey and West (1987) standard errors.

Fig. 3.17. Monetary Policy Transmission and Credit - Bank-level Responses

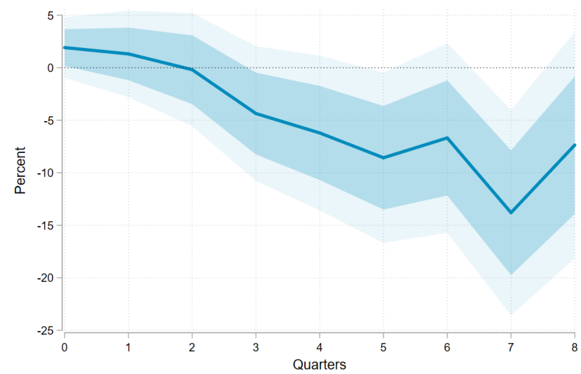
(a) **Aggregate credit**



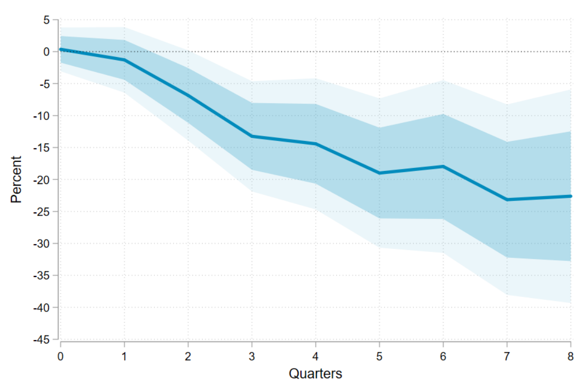
(b) **Intensive margin**



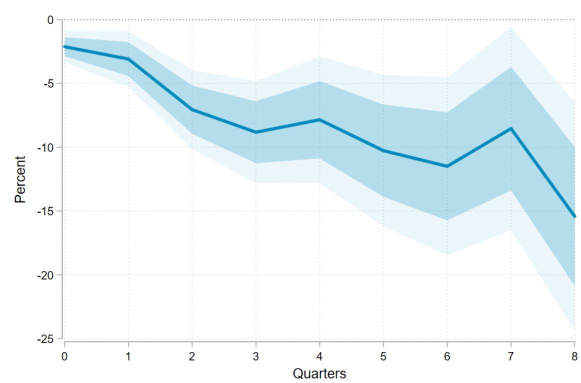
(c) **Extensive margin**



(d) **Creation**



(e) **Destruction**

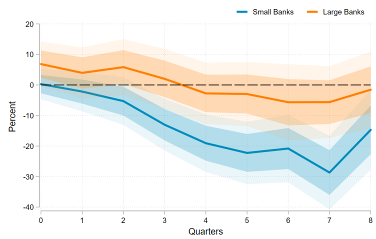


Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, and (c) extensive margin. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.17) and the "purified" monetary policy surprises from Jarociński and Karadi (2020). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock and the y-axis is in percent. The blue-shaded areas correspond to the 68% (dark blue) and 90% (light blue) confidence intervals constructed using Newey and West (1987) standard errors.

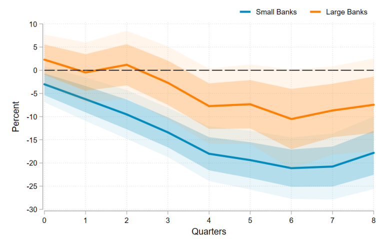
Fig. 3.18. Monetary Policy Transmission and Credit – Bank Characteristics

A. Small vs. Large Banks

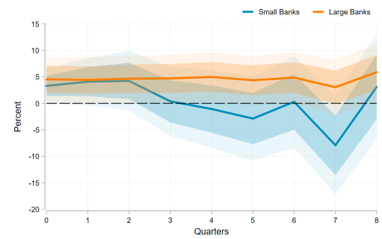
(a) **Aggregate credit**



(b) **Intensive margin**

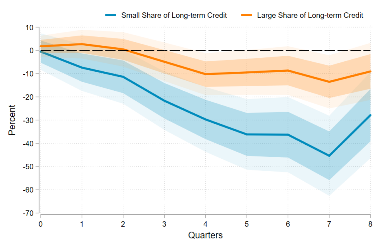


(c) **Extensive margin**

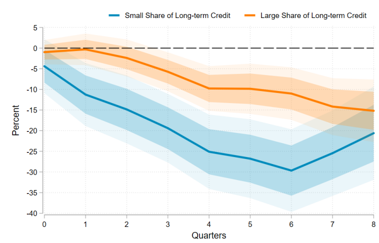


B. Small vs. Large Share of Long-term Credit

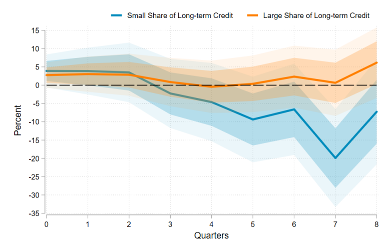
(d) **Aggregate credit**



(e) **Intensive margin**

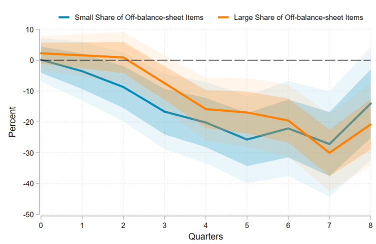


(f) **Extensive margin**

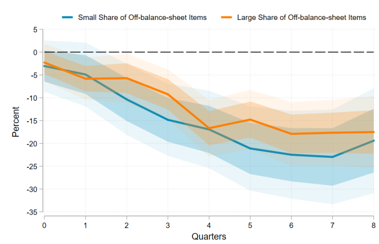


C. Small vs. Large Share of Off-Balance-Sheet Items

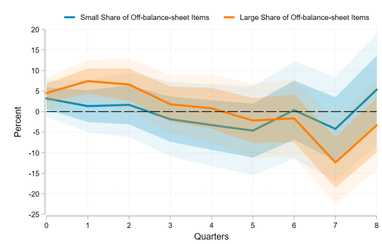
(g) **Aggregate credit**



(h) **Intensive margin**



(i) **Extensive margin**



Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, and (c) extensive margin. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.17) and the "purified" monetary policy shocks from Jarociński and Karadi (2020). The local projections are estimated separately for small vs. large banks (Panel A), banks with small vs. large share of long-term credit (Panel B) and banks with small vs. large share of off-balance-sheet credit items (Panel C). Banks are classified into small (large) groups if they are below (above) the median threshold for (i) total credit exposure, (ii) share of long-term credit, and (iii) share of off-balance-sheet items. The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The shaded areas correspond to the 68% (dark color) and 90% (light color) confidence intervals constructed using Newey and West (1987) standard errors.

Online Appendix – Not For Publication

B. Data and Variable Construction

B.1. French Credit Register (SCR)

Our raw database excludes (i) sole entrepreneurs and (ii) all firms belonging to the financial sector and public administrations. We keep only those firm-branch observations with non-missing data on firm and bank identifiers. We also remove (i) observations for bank branches located in Corsica as well as in overseas departments and territories, and (ii) branch-firm linkages for non-resident firms. We then follow standard filters for firms within our sample and delete observations for (i) various legal firm categories under French civil, commercial, or administrative law that are irrelevant for our analysis (e.g., parishes, unions, cooperatives, etc.); and (ii) financial and insurance companies, public administration, and various liberal professions. Finally, we allocate banks in our sample to a unique banking group identifier: we drop all banks that belong to nontraditional banking groups or non-credit intermediaries (e.g., public banks and financial institutions).

B.2. Balance Sheet Data (FIBEN & BRN)

We use two different datasets to gather information on French firms' balance sheets. First, FIBEN (*Fichier Bancaire des Entreprises*) accounting data are extracted from the individual company accounts. These are collected yearly through the branch network of Banque de France based on fiscal documents (i.e., balance sheet and income statements). The data collection covers all companies conducting business in France whose annual turnover exceeds EUR 0.75 million or whose bank debt exceeds EUR 0.38 million. We exploit this database to obtain relevant firm-level variables such as firm total assets, leverage, and employment. The dataset also provides information about the age of the firm, its 2-digit industry, and whether it is part of a group or a standalone company. It also contains a unique firm identifier that allows data to be merged with the SCR. Second, the BRN (*Benefices Industriels et Commerciaux - Regime Normal*) dataset is produced by the INSEE (*Institut National de la Statistique et des Etudes Economiques*) and gathers the balance sheet information of firms that opt for the *standard fiscal regime*. It provides information on employment, sales, value added, and the breakdown of investment for all firms of all sectors from 1998 through 2016.

B.3. Banking Mergers and Acquisitions (M&As)

In order to keep track of bank M&As, we rely on data from the French Supervision and Prudential Authority (ACPR). Our dataset gathers all the M&A operations involving banks located within the French territory and includes the date of the transaction as well as the identity of acquiring and acquired banks.

B.4. Public Banks

Due to their "nonstandard" objectives, we remove the following public banks from the sample:

- *Caisse nationale des Telecom* (Bank identifier: 15379)
- *Caisse nationale des autoroutes* (Bank identifier: 15389)
- *Groupe banque de development des PME* (BPI (initially titled OSEO), with bank identifiers: 10048, 13328, 13810, 13880, 14138, 18710, 19510, and 18359)
- *Groupe CDC* (Bank identifiers: 23930, 40031, 60030, and 60070)
- *Groupe credit logement* (Bank identifier: 19230)

B.5. Other Reporting Issues

The reporting methodology of the SCR has evolved constantly over the past two decades. We document here some of the issues that directly impact our tabulations and our corresponding adjustment. For example, in 2003Q3, the French Central Bank credit grading scale was amended (going from *cotation BDF* to *cotation NEC*); we use a correspondence table provided by Banque de France to ensure a consistent measure of the credit quality of borrowers. In 2012Q1, the reporting of non-performing loans was modified, which creates a minor discontinuity in some of our aggregate series. All the non-performing credit was indeed previously allocated to long-term credit (even if the maturity was shorter than one year), but after 2011Q4 its reporting was broken down into long-term and short-term categories. This evolution directly affects our measures of the number of existing relationships with short-term vs. long-term credit. We decided to artificially keep the pre-2012 norm active until the end of our sample and to re-classify relationships based on their initial maturities. Finally, and despite our efforts, we were unable to properly deal with a change in the reporting methodology for credit guarantees, occurring at the end of 2005, that led to a spike in gross flows around 2005Q4 and 2006Q1. For each gross flow time series, we manually replace this one data point at time t based on the midpoint derived from the time $t-1$ and $t+1$ data. That said, working with this data point or simply omitting it from the analysis

doesn't substantially affect any of our results.

C. Credit Relationship Flows – Additional Descriptive Results and Robustness Checks

Cross-section. Table 3.6 reports additional summary statistics in the cross-section. We note that overall there is a significant degree of heterogeneity across banks, firms, and bank-firm matches, further highlighting the importance of jointly analyzing the extensive and intensive margins of credit.[41] For example, bank size (as measured by the number of serviced borrowers or total credit exposure) is heavily skewed with a median of 77 borrowers (or equivalently EUR 137 million), with the 95^th percentile standing at over 4000 borrowers (or EUR 3.8 billion). In the same vein, relationship duration and credit exposure measures also exhibit a large degree of dispersion across relationships, with interquartile ranges spanning 5.9-24.4 quarters, and 116-429 thousand EUR, respectively.

Table 3.6: Summary Statistics: Cross-sectional Results

This table reports cross-sectional summary statistics for the period 1999Q1-2016Q4. Relationship duration is measured in quarters. All credit variables are in thousands of Euro unless specified otherwise and are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. We display the mean, 25th, 50th, 75th, 95th, and 99th percentiles for all variables over the sample period.

Percentile	p25	p50	p75	p95	p99	Mean
Number of banks per firm	1	1	1	3.1	5.4	1.36
Number of firms per bank	10.8	77.7	772.8	4,016.4	8,951.8	802.5
Bank size (EUR M)	16.4	137.2	820.9	3,839.1	12,856.6	853.1
Firm size (EUR M)	0.1	0.2	0.5	2.5	12.6	1.4
Duration	5.9	13.3	24.4	40.7	43	16.4
Credit exposure per match	116.4	196.2	429.4	2,179.8	12,753.1	1,032.9
Short-term debt per match	0	6.2	86.4	611.2	2,834.7	214.1
Long-term debt per match	13	101.4	223.9	1,103.3	5,180.2	413.9
Credit lines and guarantees per match	0	0	31.6	423.1	3,697.1	396.1

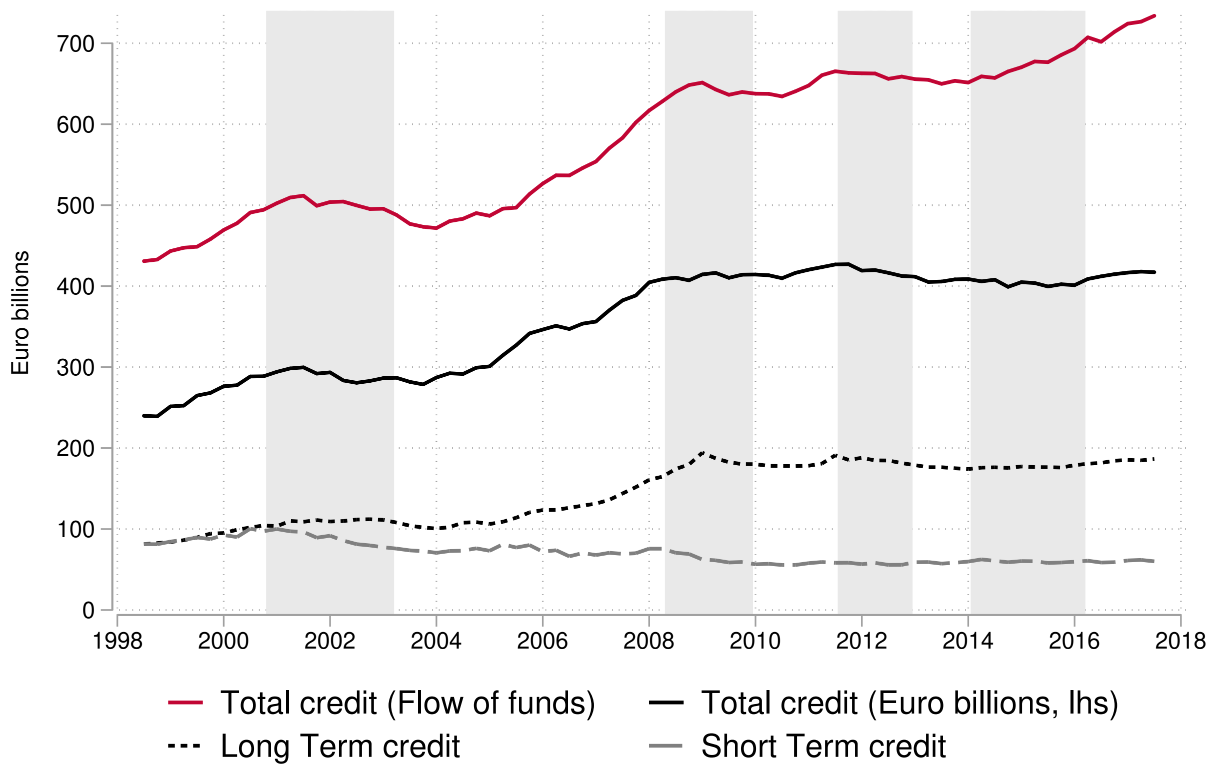
[41] On the one hand, if bank-firm matches were all identical and financial contracts were rigid (i.e., credit per match is constant throughout the relationship), then we should care only about counting the number of credit relationships in the economy (i.e., extensive margin would be a sufficient statistic for aggregate credit). On the other hand, if the processes behind the creation and destruction of bank-firm matches were frictionless and the value/quality of the relationship portable, then only the intensive margin would matter.

Table 3.7: Cyclical Properties: Lead-lag Structure

This table reports the results for cross-correlation of leads (+2 to +8) and lags (-8 to -2) for detrended credit relationship flows, relationship capital, and average credit with respect to GDP, total credit, relationship capital, and average credit, over the period 1999-2016. GDP, total credit, relationship capital, and average credit refer to the log-growth of these variables. Flow variables are detrended using an HP filter with a smoothing parameter of 1600. Nominal credit variables are deflated using the French GDP deflator, deseasonalized using the X-13 adjustment procedure, and smoothed based on MA(-1, 1).

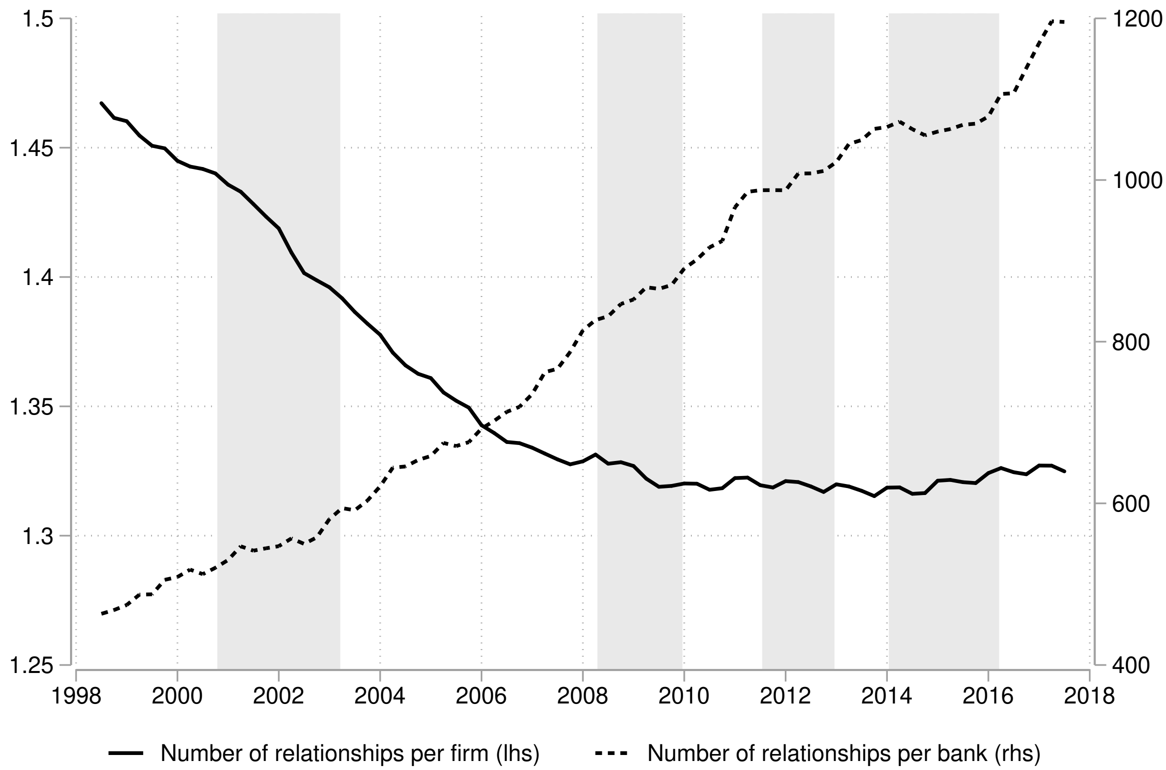
	A. Cross-correlation of GDP with:						
	x(-8)	x(-4)	x(-2)	x	x(+2)	x(+4)	x(+8)
Relationship capital	-0.36	-0.10	0.16	0.44	0.42	0.21	-0.06
Average credit	0.07	-0.04	0.16	0.23	0.32	0.35	-0.04
Creation flows	-0.29	-0.13	0.11	0.43	0.44	0.18	-0.28
Destruction flows	0.06	-0.26	-0.37	-0.26	0.14	0.40	0.21
Net flows	-0.30	-0.02	0.24	0.50	0.34	0.00	-0.34
Reallocation rate	-0.24	-0.22	-0.05	0.29	0.45	0.33	-0.16
Excess reallocation	-0.05	-0.22	-0.32	-0.14	0.19	0.37	0.16
	B. Cross-correlation of Total credit with:						
	x(-8)	x(-4)	x(-2)	x	x(+2)	x(+4)	x(+8)
Relationship capital	-0.06	0.47	0.53	0.64	0.50	0.35	0.34
Average credit	0.12	0.17	0.55	0.90	0.56	0.23	-0.10
Creation flows	0.05	0.36	0.47	0.47	0.29	-0.03	-0.10
Destruction flows	0.13	-0.13	-0.07	-0.14	0.11	0.37	-0.07
Net flows	-0.01	0.38	0.46	0.48	0.22	-0.18	-0.06
Reallocation rate	0.09	0.29	0.40	0.38	0.31	0.12	-0.12
Excess reallocation	0.08	-0.11	0.11	-0.01	0.12	0.30	-0.04
	C. Cross-correlation of Relationship capital with:						
	x(-8)	x(-4)	x(-2)	x	x(+2)	x(+4)	x(+8)
Relationship capital	0.26	0.56	0.53	1.00	0.53	0.56	0.26
Average credit	0.29	0.13	0.34	0.25	0.37	0.27	-0.22
Creation flows	0.09	0.37	0.41	0.64	0.27	0.03	-0.30
Destruction flows	0.14	0.01	-0.04	-0.26	0.12	0.11	0.03
Net flows	0.03	0.34	0.39	0.68	0.19	-0.02	-0.29
Reallocation rate	0.13	0.35	0.36	0.48	0.29	0.07	-0.27
Excess reallocation	0.06	0.10	0.17	0.01	0.25	0.10	-0.12
	D. Cross-correlation of Average credit with:						
	x(-8)	x(-4)	x(-2)	x	x(+2)	x(+4)	x(+8)
Relationship capital	-0.22	0.27	0.37	0.25	0.34	0.13	0.29
Average credit	-0.01	0.14	0.50	1.00	0.50	0.14	-0.01
Creation flows	0.01	0.24	0.35	0.24	0.21	-0.06	0.03
Destruction flows	0.08	-0.17	-0.06	-0.03	0.07	0.40	-0.10
Net flows	-0.02	0.29	0.35	0.23	0.17	-0.22	0.07
Reallocation rate	0.04	0.16	0.30	0.21	0.23	0.11	-0.01
Excess reallocation	0.07	-0.19	0.04	-0.02	0.01	0.33	0.02

Fig. 3.19. Aggregate Credit: French Credit Register (SCR) vs. Flow of Funds



Notes: This figure compares the time series of aggregate bank credit obtained from the national balance sheet items (solid red line) and aggregate credit obtained from the SCR after filters (solid black line). The black dashed curve presents the time series of aggregate long-term credit (initial maturity ≥ 1 year) while the gray dashed line represents the time series of short-term credit (initial maturity < 1 year). All nominal credit variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Gray-shaded areas correspond to recession periods.

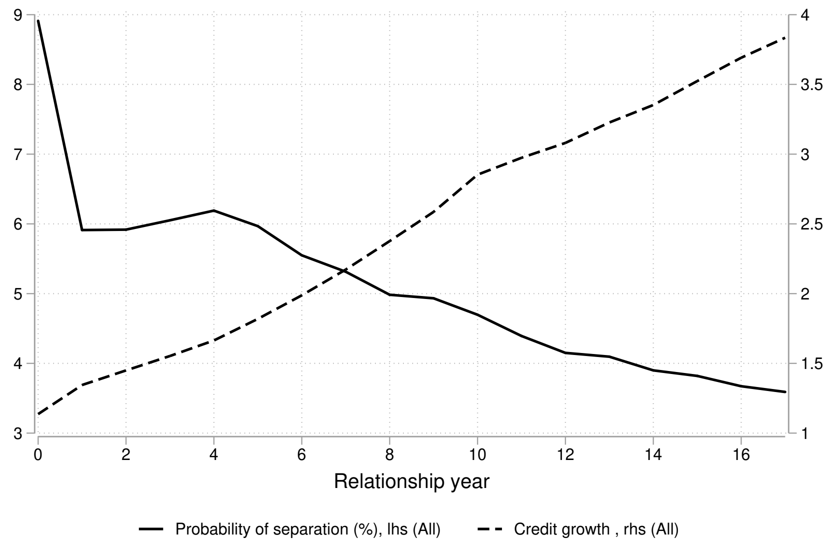
Fig. 3.20. Number of Credit Partners per Bank and per Firm



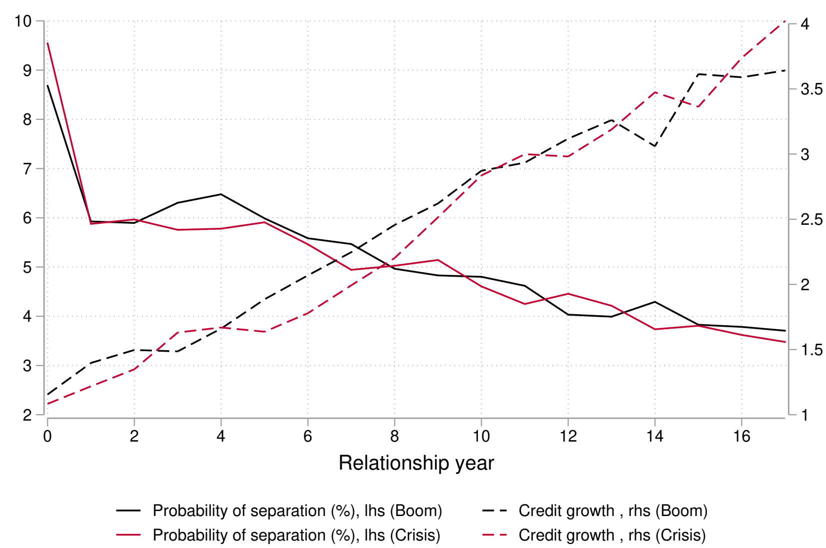
Notes: This figure reports the evolution of the number of relationships per firm (solid line) and the number of relationships per bank (dashed line) over the period 1999-2016. The sample accounts for only those relationships that are above the reporting threshold. Gray-shaded areas correspond to recession periods.

Fig. 3.21. Trajectories of Credit Growth and Separation Probability

(a) **Unconditional results**

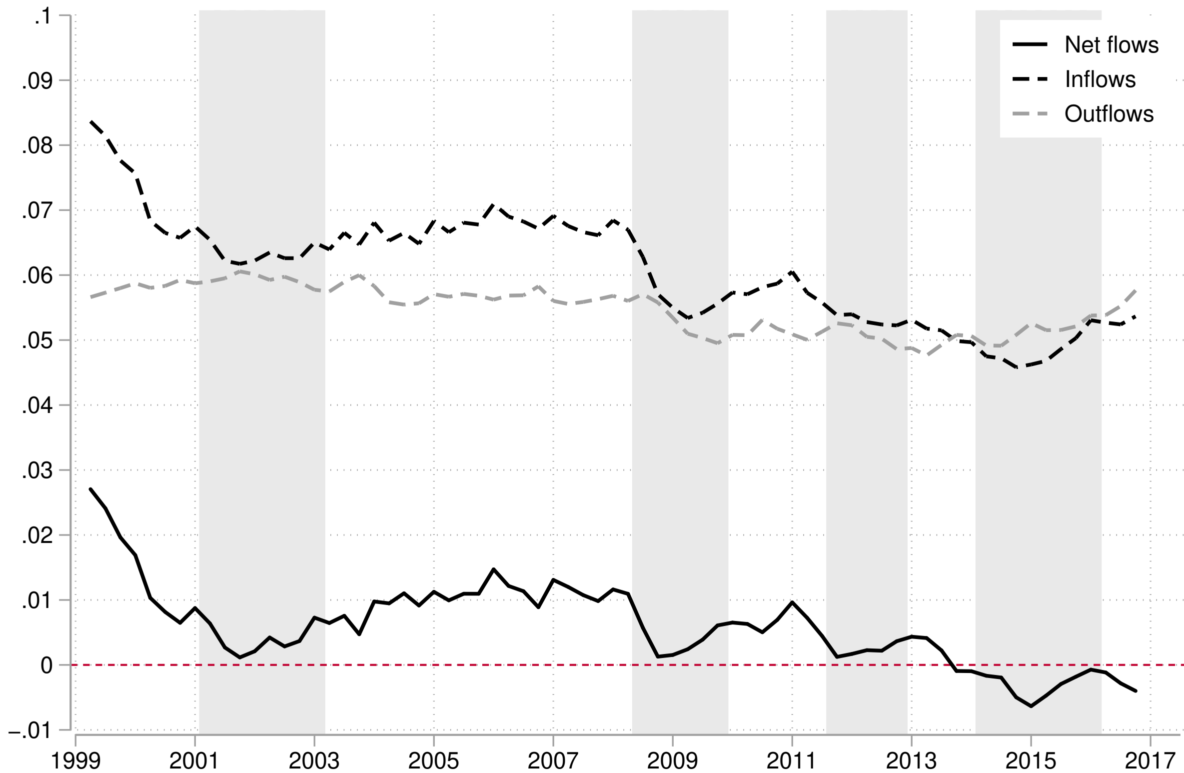


(b) **Boom vs. crisis periods**



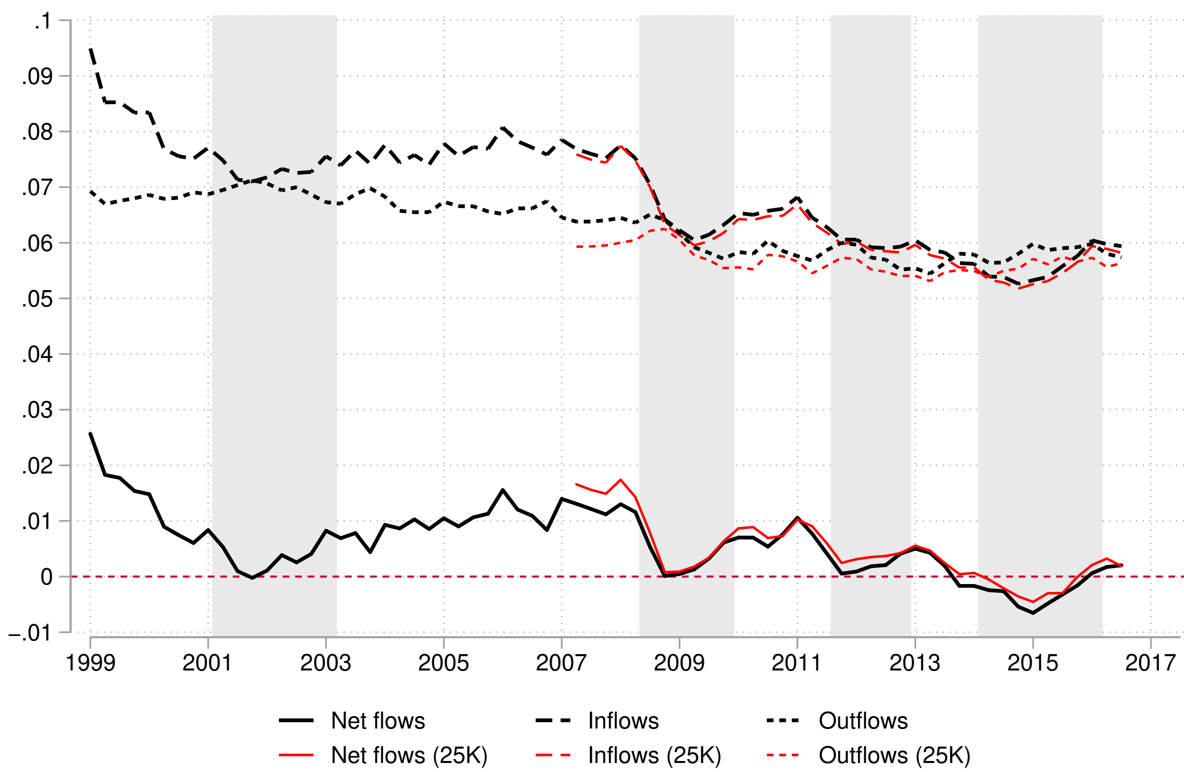
Notes: These figures show the trajectories of cross-sectional averages of credit, normalized to one at time 0 (dashed line) and separation probability (solid line) throughout the duration of a credit relationship. Panel (a) reports unconditional results, while Panel (b) reports the results for boom (in black) and crisis (in red) periods. Results are based on relationships above the reporting threshold (adjusted for inflation) and within our sample period 1999-2016.

Fig. 3.22. Credit Relationship Flows with 8-Quarter Gaps



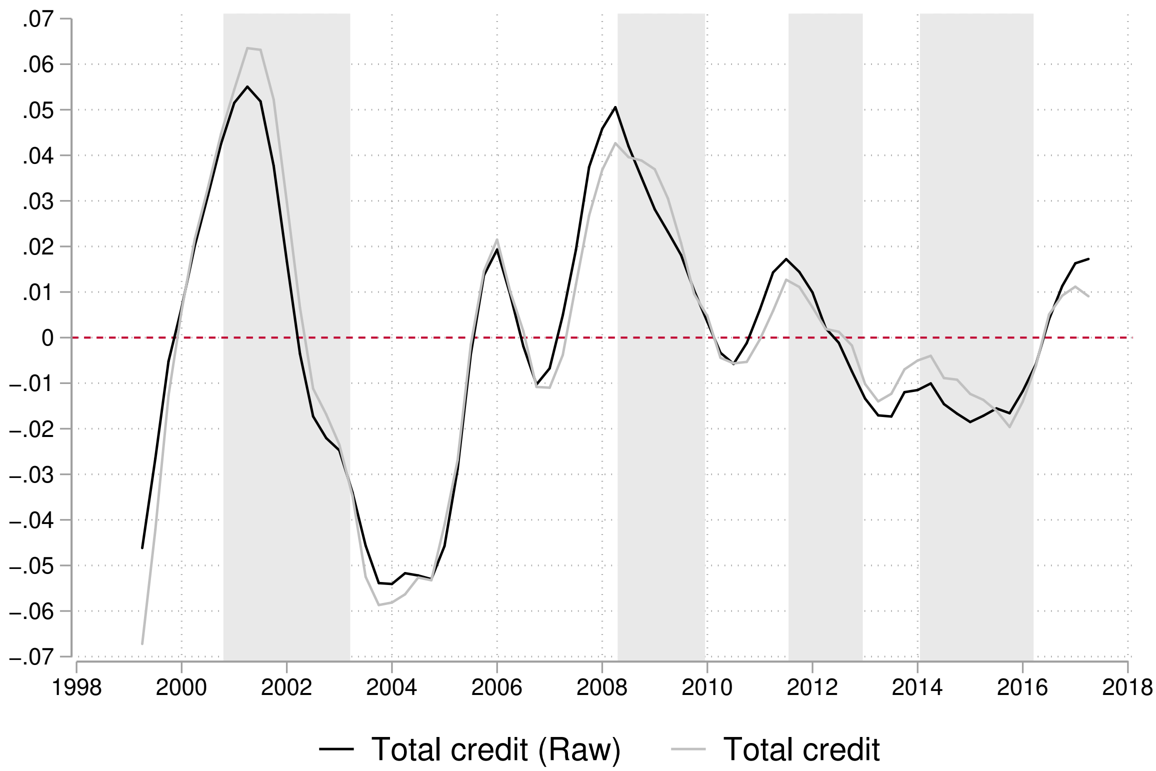
Notes: This figure shows raw net (solid black line) and gross flows of credit relationships, constructed using an 8-quarter gap. Gross creation flows (inflows) are reported in dashed black line, while gross destruction flows (outflows) are reported in dashed gray line. Results are based on relationships above the 75K Euro reporting threshold (adjusted for inflation) for the period 1999-2016. Gray-shaded areas correspond to recession periods.

Fig. 3.23. Credit Relationship Flows with the 25K Euro Threshold



Notes: This figure shows raw net (solid lines) and gross flows (creation flows in dashed lines and destruction flows in dotted lines) of credit relationships. Results are based on relationships above the 75K Euro reporting threshold for the period 1999-2016 (in black), and above the 25K Euro reporting threshold for the period 2007-2016 (in red). Both reporting thresholds are adjusted for inflation. Gray-shaded areas correspond to recession periods.

Fig. 3.24. Aggregate Credit Variations – Cyclical Components (HP Filter)



Notes: This figure shows the cyclical deviations (in log) of aggregate credit, based on the raw time series (black line), and approximated as the sum of extensive and intensive margin components from the decomposition in equation 3.10 (gray line). Nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Our sample period is 1999-2016. Gray-shaded areas correspond to recession periods.

D. Extensive/Intensive Margin Decompositions – Additional Derivations and Results

D.1. Simple Decomposition

This section provides additional derivations related to the variance decomposition of aggregate credit, based on the HP-filtered cyclical log-deviations. We start with the following identities:

$$\begin{aligned} \log(C_t) &= \log(N_t) + \log(\bar{c}_t) \\ \log(\tilde{C}_t) &= \log(\tilde{N}_t) + \log(\tilde{\bar{c}}_t). \end{aligned}$$

We can thus write:

$$\begin{aligned} \Delta \log(C_t) &= \log(C_t) - \log(\tilde{C}_t) \\ &= \Delta \log(N_t) + \Delta \log(\bar{c}_t), \end{aligned} \tag{3.18}$$

where $\tilde{X}$ is the HP-filtered trend and $\Delta X_t = X_t - \tilde{X}_t$ correspond to the cyclical deviations. We can then determine the associated betas based on this decomposition, similar to the one derived in Equations (12 - 15):

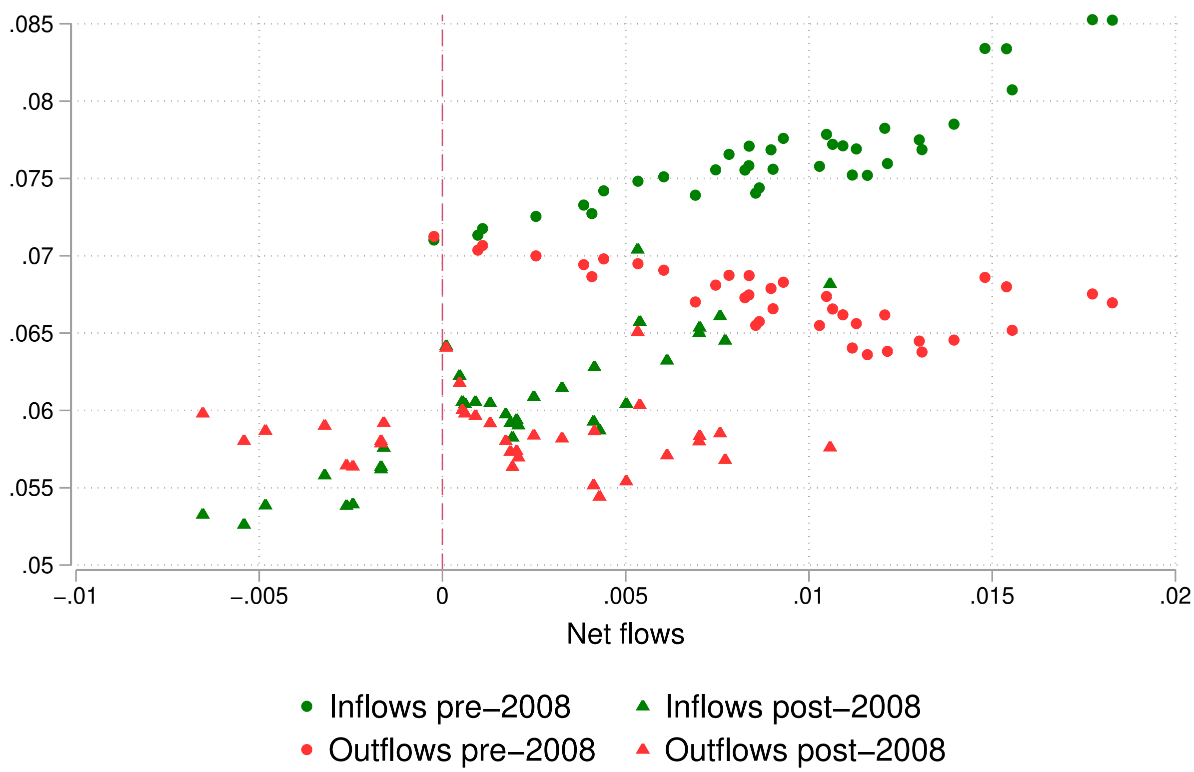
$$\begin{aligned} 1 &= \frac{\operatorname{cov}(\Delta \log(N_t), \Delta \log(C_t))}{\operatorname{var}(\Delta \log(C_t))} + \frac{\operatorname{cov}(\Delta \log(\bar{c}_t), \Delta \log(C_t))}{\operatorname{var}(\Delta \log(C_t))} \\ &= \beta_{Ext} + \beta_{Int}. \end{aligned}$$

Furthermore, we can write the following recursive expression connecting the cyclical deviations of the number of relationships to those of gross flows:

$$\begin{aligned} \Delta \log(N_{t+1}) &= \log(N_t + Pos_{t+1} - Neg_{t+1}) - \log(\tilde{N}_t + \tilde{Pos}_{t+1} - \tilde{Neg}_{t+1}) \\ &= \Delta \log(N_t) + \log(1 + c_{t+1} - d_{t+1}) - \log(1 + \tilde{c}_{t+1} - \tilde{d}_{t+1}), \end{aligned} \tag{3.19}$$

where Pos_t and Neg_t correspond to positive and negative relationship flows (in level) at time t. We can then iterate this relationship up until the time origin and rewrite the cyclical

Fig. 3.25. Creation vs. Destruction Flows: Pre- and Post-2008



Notes: This figure shows a scatter plot of the cyclical deviations (in log) of average credit and the stock of relationship, as a function of their aggregate credit counterparts. Cyclical deviations are extracted using an HP filter with a smoothing parameter of 1600. Nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Our sample period is 1998-2016.

deviations in the extensive margin as follows:

$$\begin{aligned}\Delta\log(N_{t+1}) &= \Delta\log(N_0)+\sum_{i=1}^{t+1}\log(1+c_i-d_i)-\sum_{i=1}^{t+1}\log(1+\tilde{c}_i-\tilde{d}_i)\\ &\simeq \Delta\log(N_0)+\sum_{i=1}^{t+1}(c_i-\tilde{c}_i)-\sum_{i=1}^{t+1}(d_i-\tilde{d}_i)\\ &\simeq \Delta\log(N_0)+\sum_{i=1}^{t+1}\Delta c_i-\sum_{i=1}^{t+1}\Delta d_i, \qquad (3.20)\end{aligned}$$

where the last two approximations assume small $\{c_i\}_{i=1,t+1}$ and $\{d_i\}_{i=1,t+1}$. We thus have β_{Ext} further decomposed into:

$$\beta_{Ext} \simeq \underbrace{\frac{\text{cov}(\sum_{i=1}^{t+1}\Delta c_i,\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))}}_{\beta_{Pos}}+\underbrace{\frac{\text{cov}(-\sum_{i=1}^{t+1}\Delta d_i,\Delta\log(C_t))}{\text{var}(\Delta\log(C_t))}}_{\beta_{Neg}} \qquad (3.21)$$

D.2. *Alternative Decomposition 2*

The same logic applies for alternative decompositions. We start with:

$$C_{t+1}=C_t+\underbrace{n^i_{t+1}\Delta C^i_{t+1}}_{T_{1,t+1}}+\underbrace{n^n_{t+1}\bar{C}^n_{t+1}}_{T_{2,t+1}}-\underbrace{n^s_{t+1}\bar{C}^s_{t+1}}_{T_{3,t+1}}. \qquad (3.22)$$

Assuming small $\frac{T_{1,t+1}}{C_t}$ $+\frac{T_{2,t+1}}{C_t}$ and $\frac{-T_{3,t+1}}{C_t}$, we can write:

$$\begin{aligned}\Delta\log(C_{t+1}) &= \Delta\log(C_t)+\Delta\log(1+\frac{T_{1,t+1}}{C_t}+\frac{T_{2,t+1}}{C_t}+\frac{-T_{3,t+1}}{C_t})\\ &\simeq \sum_{i=1}^{t+1}\Delta\frac{T_{1,i}}{C_{i-1}}+\sum_{i=1}^{t+1}\Delta\frac{T_{2,i}}{C_{i-1}}+\sum_{i=1}^{t+1}\Delta\frac{-T_{3,i}}{C_{i-1}}.\end{aligned}$$

Hence,

$$\begin{aligned}\text{var}(\Delta\log(C_t)) &\simeq \text{cov}(\sum_{i=1}^{t}\Delta\frac{T_{1,i}}{C_{i-1}},\Delta\log(C_t))\\ &\quad+\underbrace{\text{cov}(\sum_{i=1}^{t}\Delta\frac{T_{2,i}}{C_{i-1}},\Delta\log(C_t))}_{Entry}+\underbrace{\text{cov}(\sum_{i=1}^{t}\Delta\frac{-T_{3,i}}{C_{i-1}},\Delta\log(C_t))}_{Exit}, \qquad (3.23)\end{aligned}$$

and, after dividing each side by $\mathrm{var}(\Delta\log(C_t))$:

$$1 \simeq \beta_{Int} + \underbrace{\beta_{Entry} + \beta_{Exit}}_{\beta_{Ext}}.$$

D.3. A Third Decomposition: Gross Intensive Credit Flows (Decomposition 3)

We finally present another alternative decomposition allowing for the distinction between positive and negative (intensive) credit flows for incumbent, new, and severed relationships. This version is based on gross intensive flows, rather than on "pure" extensive vs. intensive margin. It is somewhat close to decomposition 2, although it comes with some minor adjustments. We start with the following identity:

$$C_{t+1} = C_t + Pos^i_{t+1} + Pos^n_{t+1} - Neg^i_{t+1} - Neg^s_{t+1}, \tag{3.24}$$

where Pos^i_t and Neg^i_t represent positive and negative flows of incumbent credit relationships, while Pos^n_t represents positive flows associated with new relationships, and Neg^s_t represents the negative flows associated with newly severed ones.

We can then derive the log-growth in credit as:

$$\begin{aligned}\Delta\log(C_{t+1}) &= \log(1 + \frac{Pos^i_{t+1}}{C_t} + \frac{Pos^n_{t+1}}{C_t} - \frac{Neg^i_{t+1}}{C_t} - \frac{Neg^s_{t+1}}{C_t}) \\ &\simeq 1 + \frac{Pos^i_{t+1}}{C_t} + \frac{Pos^n_{t+1}}{C_t} - \frac{Neg^i_{t+1}}{C_t} - \frac{Neg^s_{t+1}}{C_t}. \end{aligned} \tag{3.25}$$

And, similar to previous decompositions, we get:

$$1 \simeq \beta_{\mathrm{Pos}^i} + \beta_{\mathrm{Neg}^i} + \beta_{\mathrm{Pos}^n} + \beta_{\mathrm{Neg}^s}.$$

For the HP filter approach, we can equivalently write:

$$C_{t+1} = C_t + Pos^i_{t+1} + Pos^n_{t+1} - Neg^i_{t+1} - Neg^s_{t+1}. \tag{3.26}$$

and, assuming small $\frac{Pos^i_{t+1}}{C_t}$, $\frac{Pos^s_{t+1}}{C_t}$, $\frac{Neg^i_{t+1}}{C_t}$, and $\frac{Neg^s_{t+1}}{C_t}$,

$$\begin{aligned}
\Delta\log(C_{t+1}) &= \Delta\log(C_t)+\Delta\log(1+\frac{Pos^i_{t+1}}{C_t}+\frac{Pos^n_{t+1}}{C_t}-\frac{Neg^i_{t+1}}{C_t}-\frac{Neg^s_{t+1}}{C_t}) \\
&\simeq \Delta\log(C_0)+\sum_{i=1}^{t+1}\Delta\frac{Pos^i_i}{C_{i-1}}+\sum_{i=1}^{t+1}\Delta\frac{Pos^n_i}{C_{i-1}}-\sum_{i=1}^{t+1}\Delta\frac{Neg^i_i}{C_{i-1}}-\sum_{i=1}^{t+1}\Delta\frac{Neg^s_i}{C_{i-1}} \qquad (3.27)
\end{aligned}$$

We can eventually derive the variance decomposition as:

$$\begin{aligned}
\mathrm{var}(\Delta\log(C_t)) &= \mathrm{cov}(\sum_{i=1}^{t}\Delta\frac{Pos^i_i}{C_{i-1}},\Delta\log(C_t))+\mathrm{cov}(-\sum_{i=1}^{t}\Delta\frac{Neg^i_i}{C_{i-1}},\Delta\log(C_t)) \\
&\quad +\mathrm{cov}(\sum_{i=1}^{t}\Delta\frac{Pos^n_i}{C_{i-1}},\Delta\log(C_t))+\mathrm{cov}(\sum_{i=1}^{t}-\Delta\frac{Neg^s_i}{C_{i-1}},\Delta\log(C_t)), \qquad (3.28)
\end{aligned}$$

and write after dividing each side by $\mathrm{var}(\Delta\log(C_t))$:

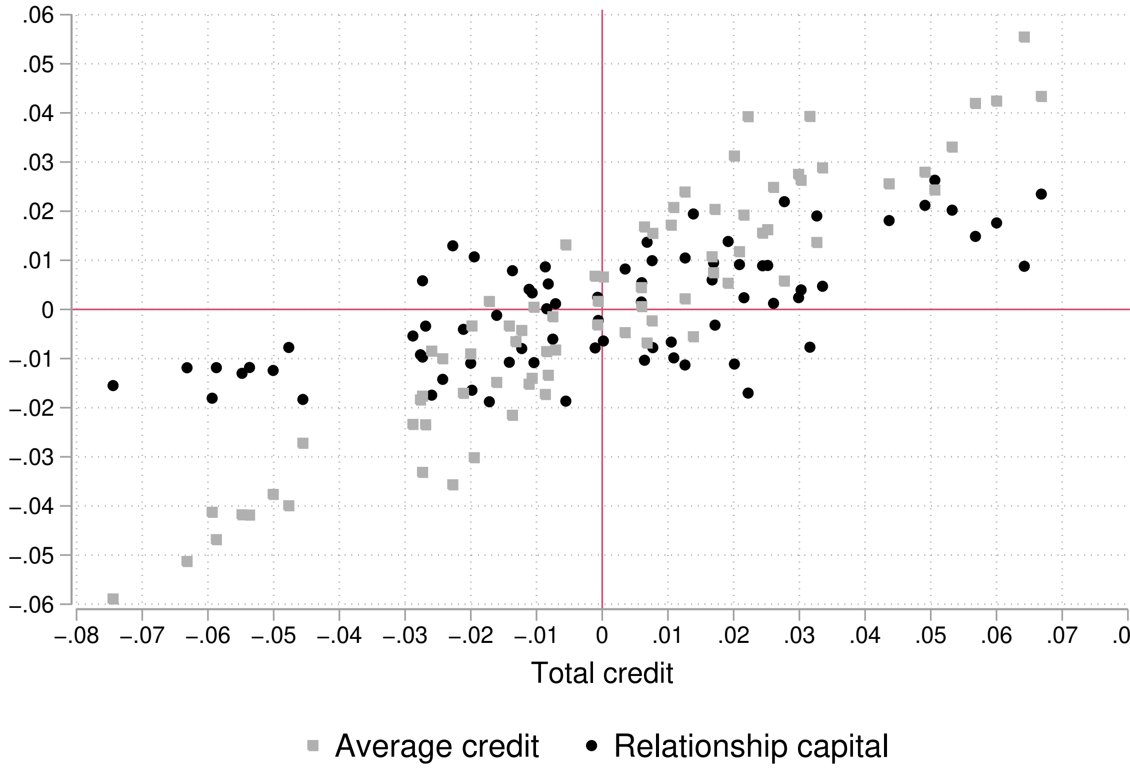
$$1 \simeq \underbrace{\beta_{Pos_i}+\beta_{Neg_i}}_{\beta_{Int}}+\underbrace{\beta_{Pos_n}+\beta_{Neg_s}}_{\beta_{Ext}}.$$

Table 3.8: Variance Decomposition: Intensive vs. Extensive Margins (Decomposition 3)

This table reports the results for variance decompositions of aggregate credit fluctuations over the period 1999-2016. The intensive/extensive margin decompositions are derived based on first-differences and log-deviations from trend obtained from HP filter with a smoothing parameter of 1600. All nominal credit variables are deflated using the French GDP deflator, deseasonalized using the X-13 seasonal adjustment procedure, and smoothed based on MA(-1, 1).

Decomposition 3				
First-Difference	Intensive Margin		Extensive Margin	
	0.52		0.48	
	Pos. flows - Incumbent	Neg. flows - Incumbent	New bank-firm effect	Severed bank-firm effect
	0.79	-0.27	0.74	-0.26
HP Filter	Intensive Margin		Extensive Margin	
	0.61		0.42	
	Pos. flows - Incumbent	Neg. flows - Incumbent	New bank-firm effect	Severed bank-firm effect
	0.53	0.08	0.72	- 0.30

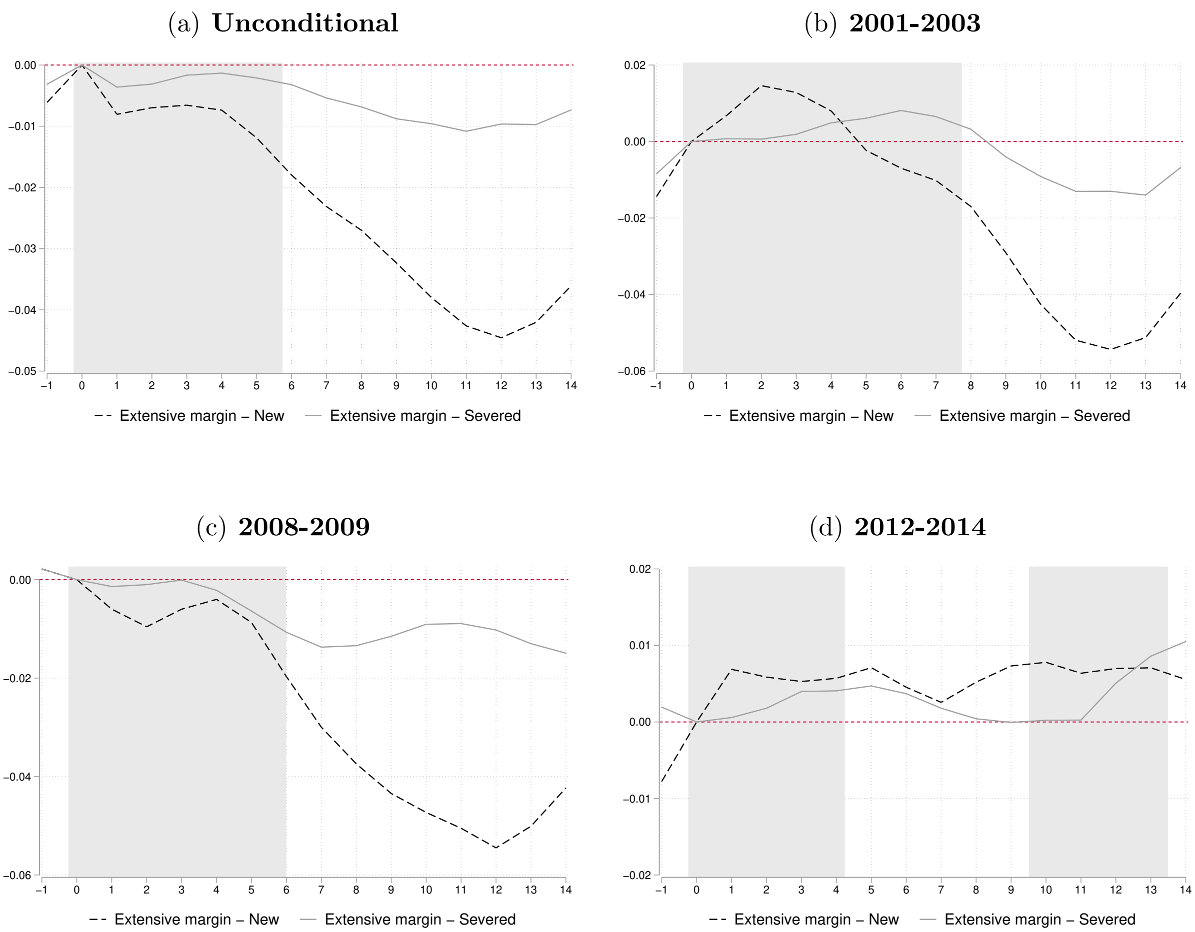
Fig. 3.26. Extensive vs. Intensive Margins: Cyclical Deviations



Notes: This figure shows a scatter plot of the cyclical deviations (in log) of average credit and the stock of relationship, as a function of their aggregate credit counterparts. Cyclical deviations are extracted using an HP filter with a smoothing parameter of 1600. Nominal variables are deflated using the French seasonally adjusted GDP Implicit Price Deflator obtained from the FRED database. Results are based on relationships above the reporting threshold (adjusted for inflation). Time series of relationship flows are first deseasonalized using the X-13 seasonal adjustment procedure and smoothed out using MA(-1,1). Our sample period is 1999-2016.

E. Anatomy of a Crisis – Additional Figures

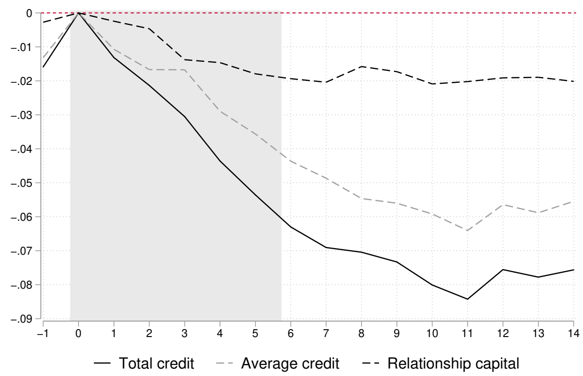
Fig. 3.27. Anatomy of a Crisis – Decomposition 2 – Creation vs. Destruction



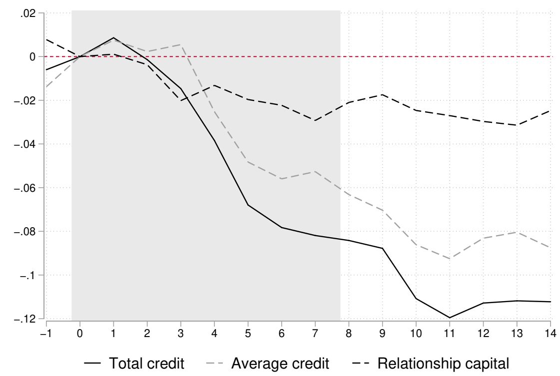
Notes: These figures report the unconditional and crisis dynamics of the creation (new) and destruction (severed) components of the extensive margin over fourteen quarters following the onset of a recession. The extensive margin is based on decomposition 2, specified in equation (3.10). All variables are normalized to 0 based on the timing of the pre-recession peak for aggregate credit, and reported in terms of log-deviations from their corresponding HP trend obtained with a smoothing parameter of 1600. Gray-shaded areas correspond to the recession period.

Fig. 3.28. Anatomy of a Crisis – Decomposition 1

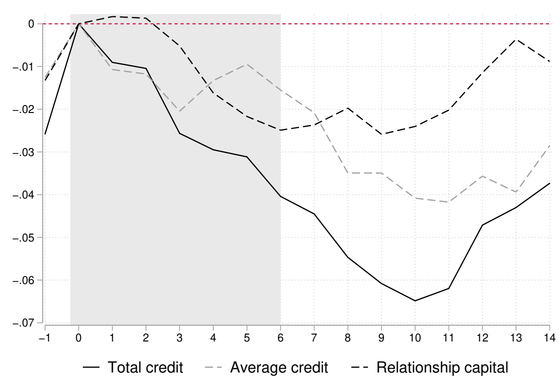
(a) **Unconditional**



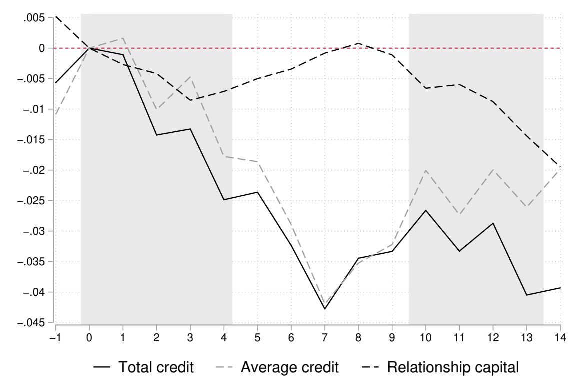
(b) **2001-2003**



(c) **2008-2009**



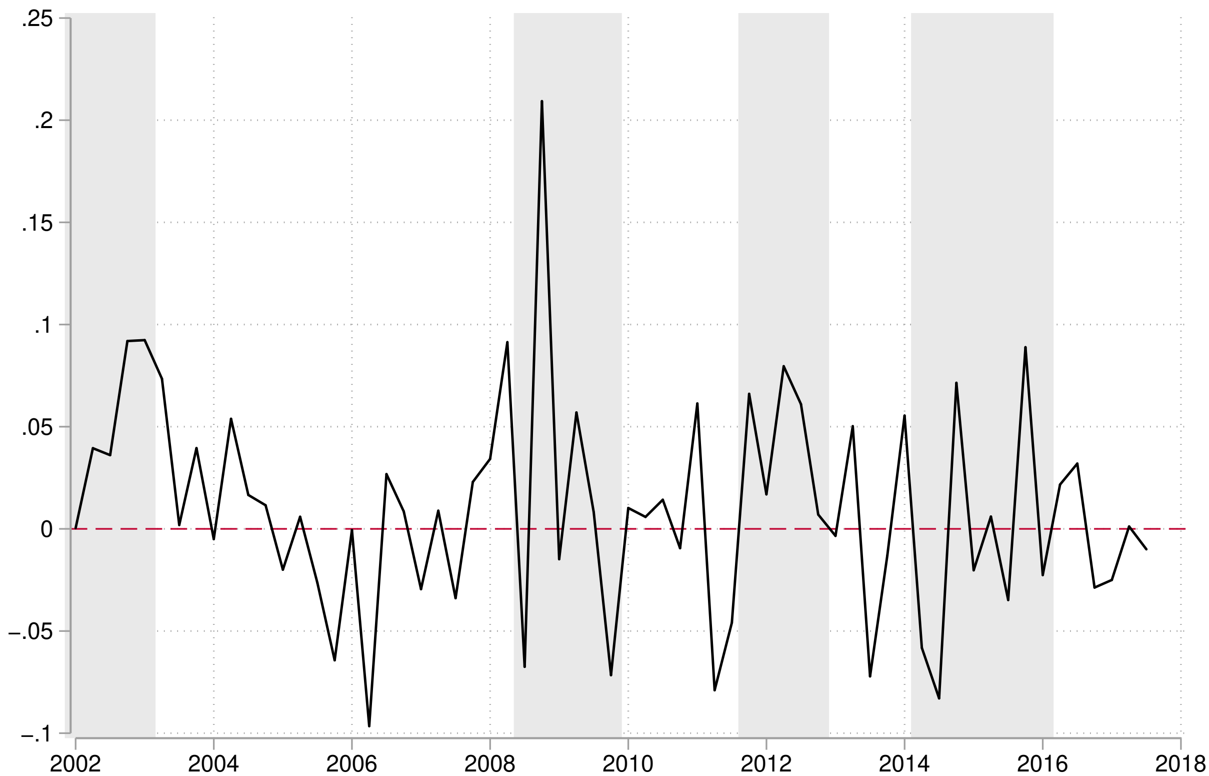
(d) **2012-2014**



Notes: These figures report the evolution of aggregate credit, average credit, and relationship capital over the fourteen quarters following the onset of each recession. Panel (a) reports unconditional results, while Panels (b), (c), and (d) report individual recessions. Due to their proximity, the recessions of 2012-2013 and 2014-2016 are shown combined in panels (d). All variables are normalized to 0 based on the timing of the pre-recession peak for aggregate credit, and reported in terms of log-deviations from their corresponding HP trend obtained with a smoothing parameter 1600. Gray-shaded areas correspond to recession periods.

F. Local Projections – Additional Figures and Results

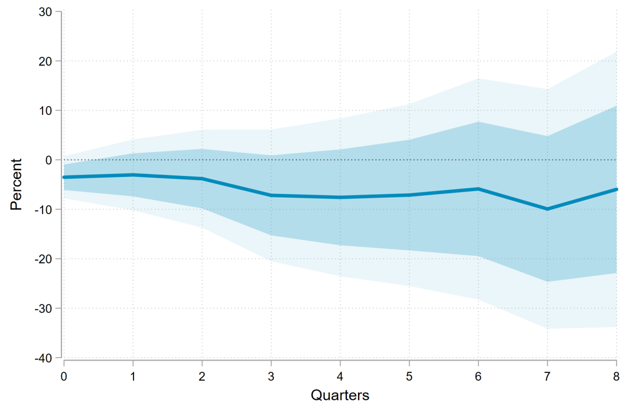
Fig. 3.29. Monetary Policy Shocks – 2002-2018



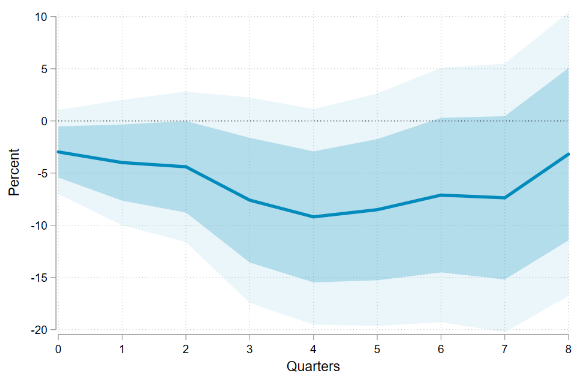
Note: These times series represent the monetary policy shocks based on the "purified" monetary policy surprises from Jarociński and Karadi (2020), aggregated at quarterly frequency over the period 2002-2018.

Fig. 3.30. Monetary Policy Transmission and Credit – Specification with Lags

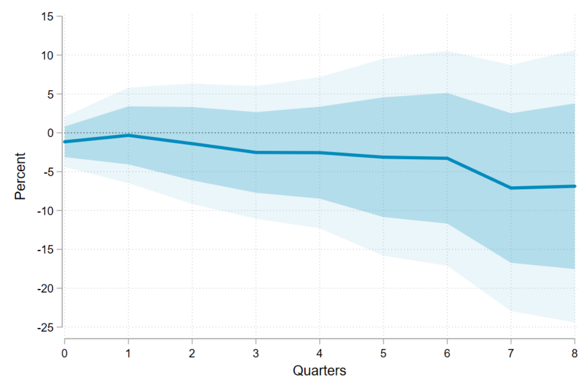
(a) **Aggregate credit**



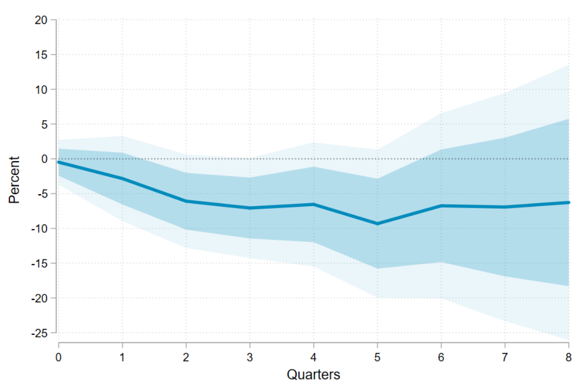
(b) **Intensive margin**



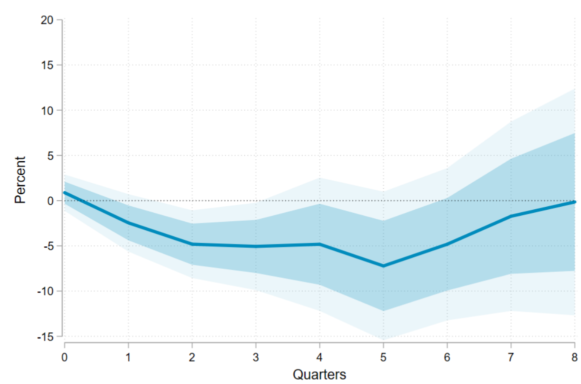
(c) **Extensive margin**



(d) **Creation**

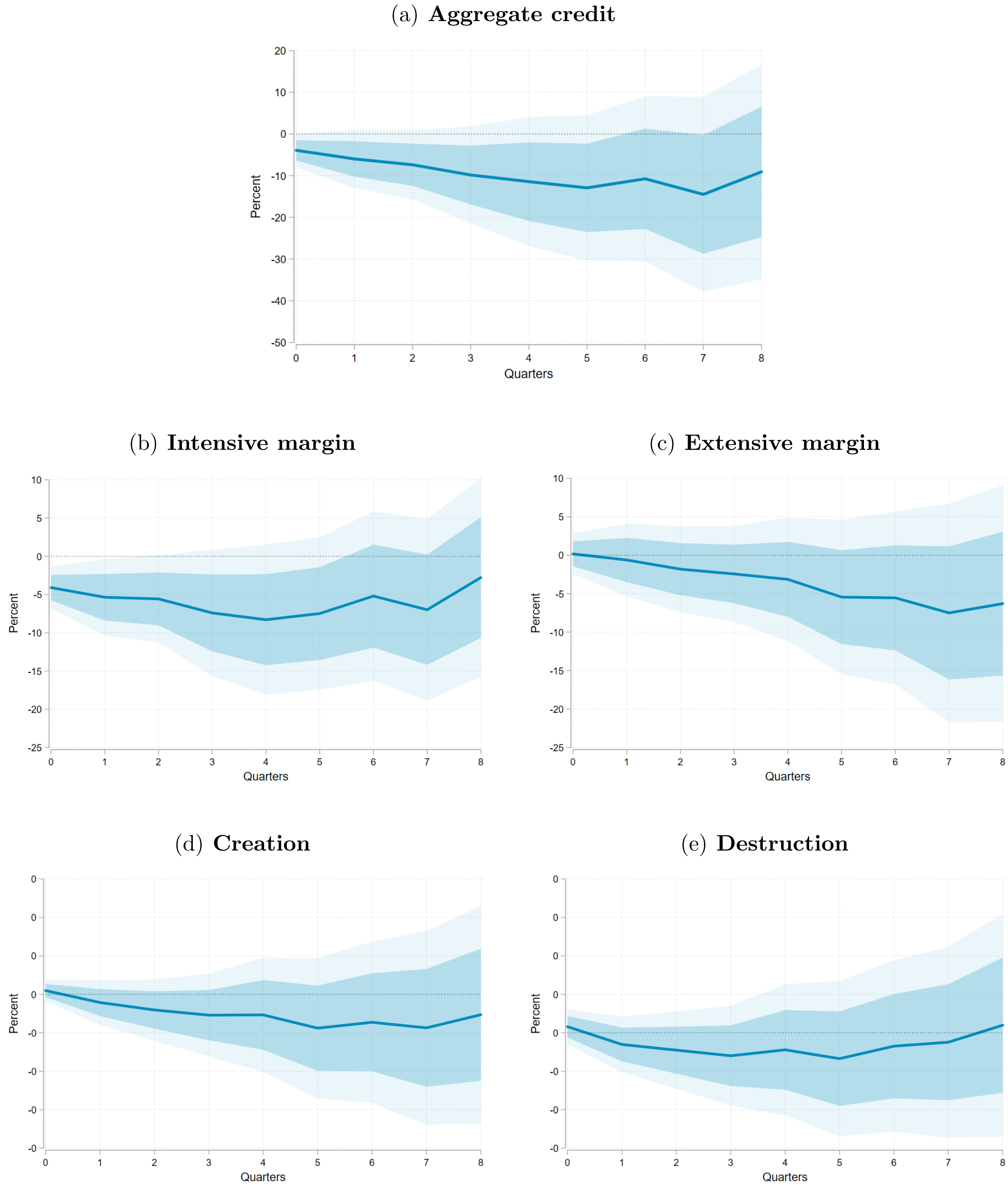


(e) **Destruction**



Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, (c) extensive margin, and the corresponding (d) creation and (e) destruction components. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.15) and the "purified" monetary policy shocks from Jarociński and Karadi (2020). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The blue-shaded areas correspond to the68% (dark blue) and 90% (light blue) confidence intervals constructed using Newey and West (1987) standard errors.

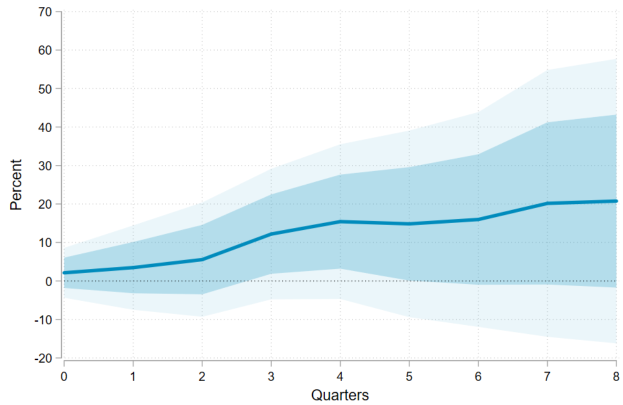
Fig. 3.31. Monetary Policy Transmission and Credit – Alternative Monetary Shocks



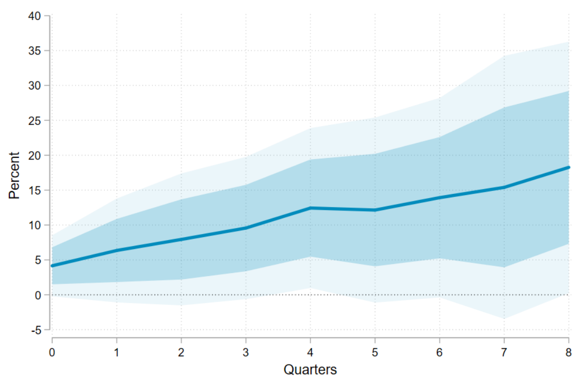
Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, (c) extensive margin, and the corresponding decomposition into (d) creation and (e) destruction components. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.15) and the "purified" monetary policy shocks from Kerssenfischer (2019). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The blue-shaded areas correspond to the 68% (dark blue) and 90% (light blue) confidence intervals constructed using Newey and West (1987) standard errors.

Fig. 3.32. Monetary Policy Transmission and Credit - ECB Information Shocks

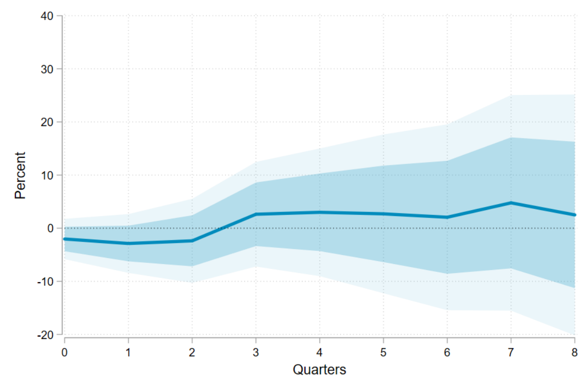
(a) **Aggregate credit**



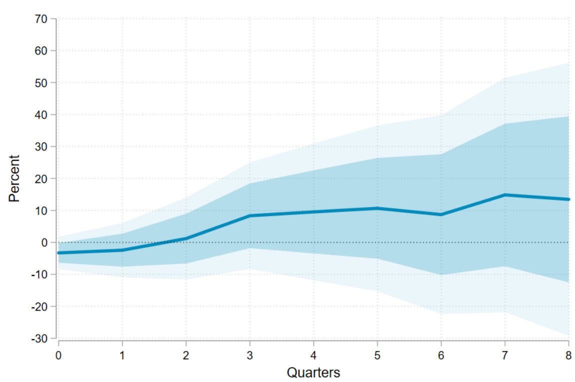
(b) **Intensive margin**



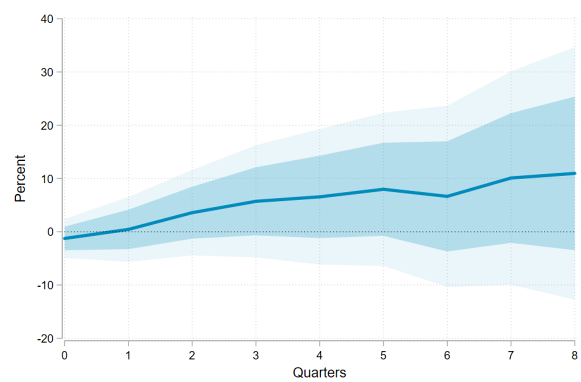
(c) **Extensive margin**



(d) **Creation**



(e) **Destruction**



Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, (c) extensive margin, and the corresponding (d) creation and (e) destruction components. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.15) and the central bank information shocks from Jarociński and Karadi (2020). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The blue-shaded areas correspond to the 68% (dark blue) and 90% (light blue) confidence intervals constructed using Newey and West (1987) standard errors.

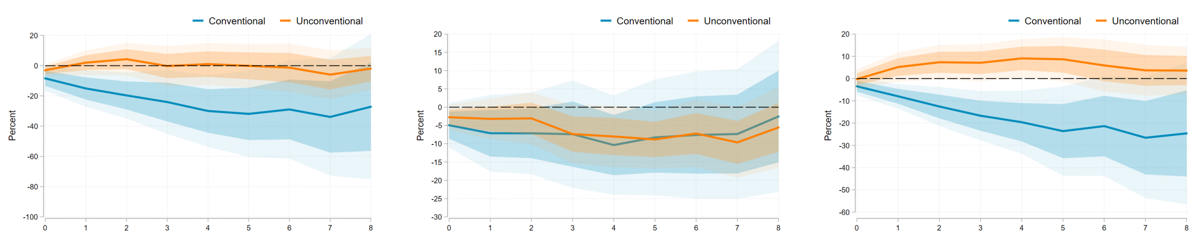
F.1. Conventional vs. Unconventional Monetary Policy

Given the significant shift in monetary policy conduct over the past two decades, the information conveyed by policy meetings has also evolved to include details about unconventional monetary policy, with potential implications for both quantities and prices. Among others, Long-Term Refinance Operations (LTRO) have been first announced in August 2007 and used extensively throughout the financial crisis and the European sovereign debt crisis. Here, we focus on the effect of monetary policy surprises when Long-Term LTRO announcements occur simultaneously. We use the event database for the ECB as collected by Cieslak and Schrimpf (2019) to determine whether a policy meeting announcement makes a reference to LTRO. We then run the specification (3.16), with a dummy variable δ_T, which equals 1 for announcements referring to LTRO and 0 for those that do not.

Our results (Figure 3.33) show that monetary policy surprises, when combined with LTRO announcements, do not generate a significant response for aggregate credit. In fact, the extensive and intensive margins appear to react in opposite, as a tightening surprise generates a decline in the intensive margin but an increase in the extensive margin. While a formal investigation of LTRO shocks is outside the scope of this paper, these results would simply argue that (tightening) monetary policy surprises tend to have more significant effects on credit when not confounded with simultaneous LTRO shocks.

Fig. 3.33. Monetary Policy Transmission – Conventional vs. Unconventional

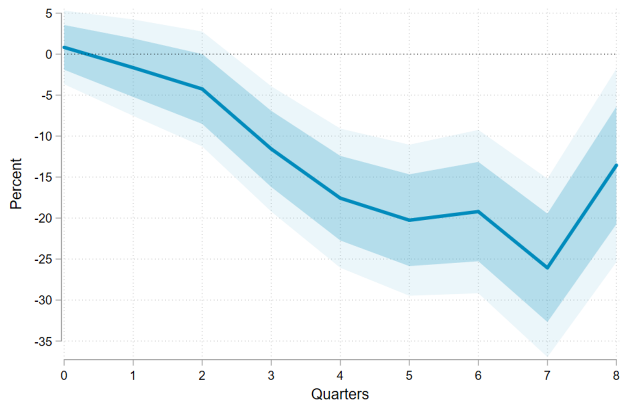
(a) **Aggregate credit** (b) **Intensive margin** (c) **Extensive margin**



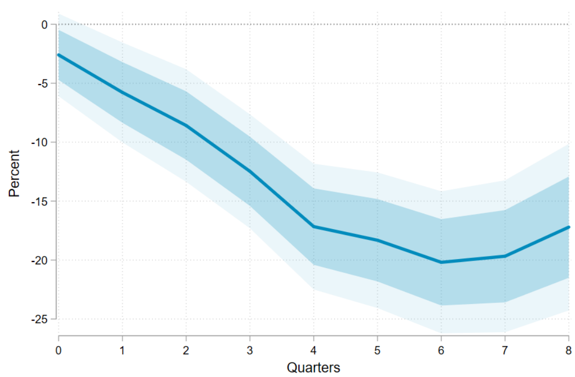
Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, and (c) extensive margin. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.15) and the "purified" monetary policy surprises from Jarociński and Karadi (2020). The sample period is 2002-2018. The local projections are estimated separately for the pre- and post-2008 periods. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The shaded areas correspond to the 68% (dark color) and 90% (light color) confidence intervals constructed using Newey and West (1987) standard errors.

Fig. 3.34. Monetary Policy Transmission and Credit – micro-level Responses (with Bank-fixed Effects)

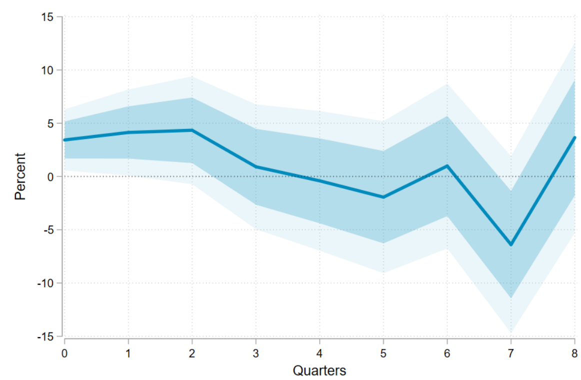
(a) **Aggregate credit**



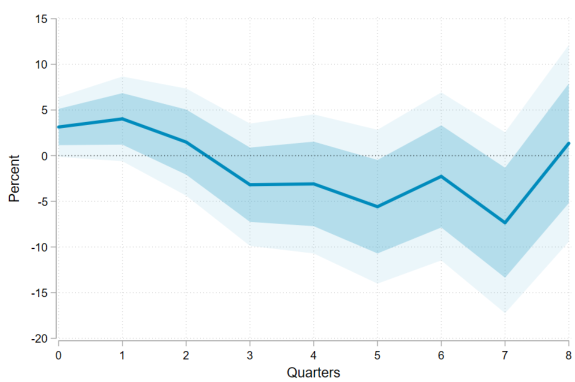
(b) **Intensive margin**



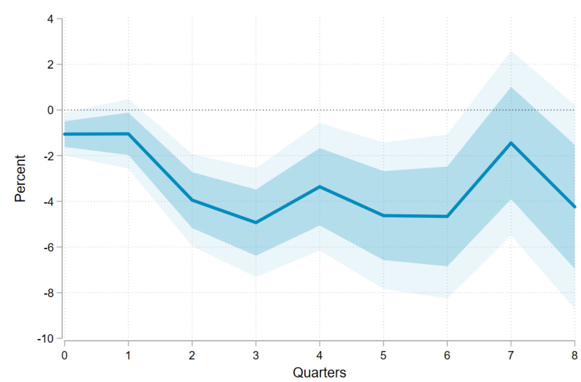
(c) **Extensive margin**



(d) **Creation**



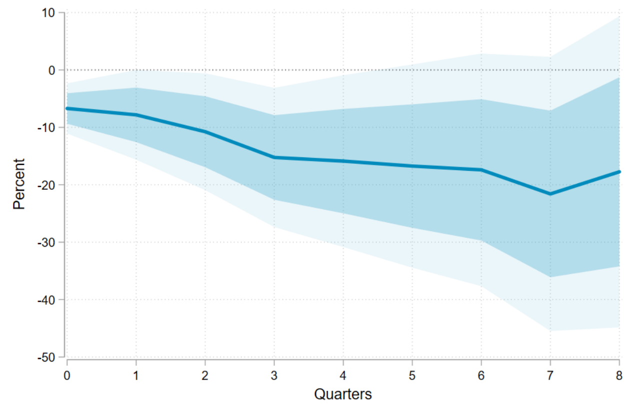
(e) **Destruction**



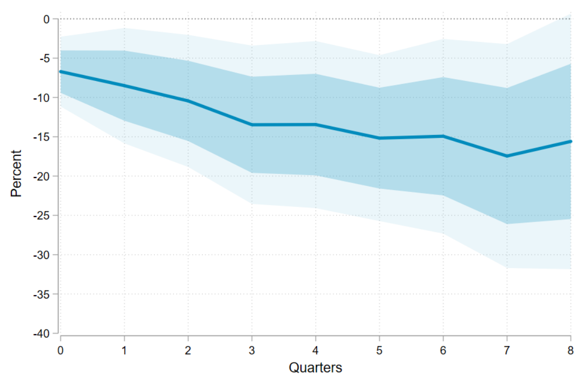
Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, (c) extensive margin, and the corresponding (d) creation and (e) destruction components. The results rely on the refined credit decomposition 2 with the local projection specification described in equation (3.17) with bank-fixed effects and the "purified" monetary policy shocks from Jarociński and Karadi (2020). The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The blue-shaded areas correspond to the 68% (dark blue) and 90% (light blue) confidence intervals constructed using Newey and West (1987) standard errors.

Fig. 3.35. Monetary Policy Transmission and Credit – Specification with Credit Decomposition 1

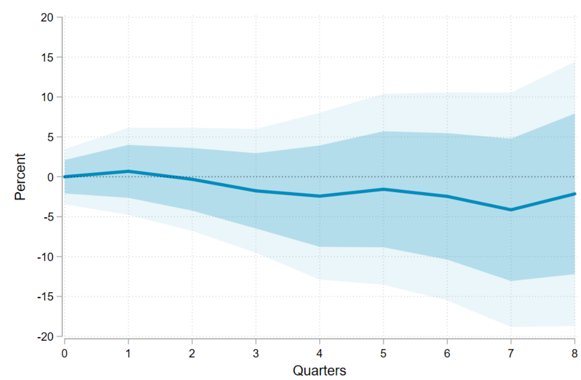
(a) **Aggregate credit**



(b) **Intensive margin**



(c) **Extensive margin**



Notes: These figures illustrate impulse responses to a one percentage point contractionary monetary policy shock for (a) aggregate credit, (b) intensive margin, and (c) extensive margin. The results rely on the simple credit decomposition 1 with the local projection specification described in equation (3.15) and the "pure" monetary policy shocks from Jarociński and Karadi (2020).The sample period is 2002-2018. The x-axis represents the number of quarters after the shock, and the y-axis is in percent. The blue-shaded areas correspond to the 90% (light blue) and 68% (dark blue) confidence intervals constructed using Newey and West (1987) standard errors.

4

Technology-induced Trade Shocks?

This chapter is based on a paper co-authored with Clément Malgouyres (IPP, PSE) and Thierry Mayer (Sciences Po, CEPR).

Abstract

In this paper, we document the presence of "technology-induced" trade in France between 1997 and 2007 and assess its impact on consumer welfare. We use the staggered roll-out of broadband internet to estimate its causal effect on the importing behavior of affected firms. Using an event-study design, we find that broadband expansion increases firm-level imports by around 25%. We further develop a model where firms optimize over their import strategy and which yields a sufficient statistics formula for the quantification of the effects of broadband on consumer welfare. Interpreted within this model, our reduced-form estimates imply that broadband internet reduced the consumer price index by 1.85% and that the import-channel, i.e. the enhanced access to foreign goods that is allowed by broadband, accounts for about 40% of that effect.

JEL codes: F14, F15, F61, F66, L23, O33
Keywords: Internet; Trade; Imports; Consumer Welfare

1. Introduction

FROM 1995 to 2008, the value of imports by high-income countries has grown twice as fast as global GDP.[1] This acceleration of globalization has induced well-documented labor market impacts (Autor et al., 2016b, summarize the recent literature on the impact of the "China shock" on labor market outcomes), as well as rises in consumer welfare through lower prices and gains in varieties (see Feenstra and Weinstein, 2017, for a recent illustration). This period was also characterized by radical innovations in information and

[1]The World Bank World Development Indicators report that the ratio of imports over GDP for high-income countries has grown from 43% to 61%.

communication technologies (ICT) and by their rapid diffusion throughout the world economy. It is most likely that the "ICT revolution" (Cohen et al., 2004) lowered the cost of carrying out international transactions and contributed to raising the pace of economic integration.[2] To the extent that ICT facilitated international trade, part of the consumer gains induced by trade development should be attributed to the diffusion of ICTs. In this paper, we test this proposition by estimating the effect of the diffusion of broadband internet on the importing behavior of French firms from 1997 to 2007 and by developing a theoretical framework to assess the impact of broadband-induced imports on consumer welfare.

Identifying the causal effect of technology on trade is generically difficult because of its endogeneity. The French data and context allow us to make progress on the causal identification of how technology affects firm-level import behavior. In terms of data, we assemble a novel dataset on broadband internet availability at the municipality level over the 1997-2007 period and combine it with information regarding firms' importing behavior. Regarding the context, we exploit the gradual roll-out of broadband internet in France, which was staggered over several years due to limited funding and completed primarily in order to maximize population coverage with only limited attention paid to local economic conditions. This setup provides natural ground for an event-study identifying how ICT availability affected firms' importing behavior.

We find that the local access to broadband internet leads to a surge in the total value of firm-level imports. Our point estimate implies a 25% increase after five years. When applied to our estimating sample and taking into account dynamic effects, the aggregate effect of broadband expansion on the value of imports in constant dollars over the 1997-2007 period was to increase its growth rate from 75% to 91%, i.e. 16 p.p. or 21%. Our results are robust to several potential threats to identification. First, we find no evidence of pre-expansion differential trends in outcomes. Second, while it is possible that broadband introduction was systematically associated with contemporaneous local economic shocks, we show that adding a rich set of city-level controls for local industry and income dynamics hardly affects our estimates. Additionally, flexibly controlling for changes in local labor market conditions, by including a large set of local fixed effects interacted with year dummies, barely changes our estimates.

We further document changes in importing activities along several margins. We find that the increase in the overall value of imports is primarily associated with an increase in the number of flows and find no effect on the average value per flow (where a flow is defined

[2]Development in ICTs, such as internet and cell phones, has also contributed to unifiying domestic markets, in particular in developing economies (Allen, 2014).

as the combination of an importing firm, an origin country and a specific product). All types of goods (intermediary, consumption and capital) are affected. We also find that broadband internet has a positive impact on firm performance as measured by value-added and sales. Importantly, the import-intensity of firms increases: the ratio of imports over sales is positively affected as well as the share of foreign inputs in overall consumption of intermediates.

Our paper focuses on the importing behavior of firms, an outcome that has been relatively underexplored in comparison to exports (see e.g. Hjort and Poulsen, 2018). This focus is motivated empirically by the fact that imports have been more dynamic than exports over the period and conceptually by the notion that imports matter most directly for consumer. Nevertheless, we present some empirical results on the export side. We find that broadband stimulated exports albeit to a lower extent than import.

Our main results rely on regressions carried out at the city level. This level of aggregation matches the level at which we have variation in exposure to treatment and allows us to capture several margins (intensive margin, extensive margin—firms starting to import—, entry of new firms) which would be missed by a firm-level panel analysis. Moreover, the model we use in the last of the paper delivers naturally a city-level estimable equation for sales and imports. We present firm-level estimates when exploring the heterogeneity of the effect along a number of dimensions, some of which intrinsically defined at the firm-level (size, sector etc.).

In the final part of the paper, we assess the welfare implications of our empirical findings through the lens of a simple but general theoretical model of firm-level imports. We consider a small open economy made of up of several cities. Firms vary idiosyncratically in terms of productivity across cities and choose the cities that maximize their expected profits (Suárez Serrato and Zidar, 2016; Fajgelbaum et al., 2019). We use the standard monopolistic competition cum CES demand setup for final goods so as to link firm-level sales with (quality-adjusted) unit cost. On the production side, hinging upon Blaum et al. (2018), firms combine labor with domestic and imported inputs. Our model features the sufficiency result highlighted by Blaum et al. (2018): the firm-level domestic share of inputs fully characterizes the contribution of imports of intermediates to the reduction of its unit cost. The generality of the setup allows us to remain agnostic regarding which type of trade costs (variable, fixed per destination or product, search friction etc.) is affected by the broadband internet expansion shock. It also generates a very parsimonious framework for welfare analysis. The overall effect of access to fast internet on the consumer price index and the contribution of enhanced access to foreign inputs to that overall effect (which we refer to as the *import*

channel) are expressed as a function of two reduced-form estimates and three parameters to be calibrated (either from descriptive statistics or from the existing relevant literature). Under our preferred values for calibrated parameters, our event-study estimates imply that broadband internet led to a price index decrease of 1.85%. The import channel contributed up to 0.75%, i.e. about 40% of the overall effect. We complete this analysis with an extension that considers the impact of broadband on exports and nominal wages. Results from the extended welfare analysis indicate that the *export channel* played a much more limited role than the import one in driving welfare gains from broadband expansion.

The remainder of the paper is structured as follows. We start by relating our paper to several streams of papers that have studied the links between technology improvements and trade. We then present the data and institutional context in Section 3. We detail our empirical approach in Section 4. Baseline results and robustness checks are presented in Section 5. Section 6 analyses the heterogeneity of the effect by types of goods and sourcing country. The section additionally presents findings on firm sales, value-added, exports as well as import intensity (imports over sales and imports over overall intermediate consumption) that are important inputs into the welfare analysis. The section presents results from an heterogeneity analysis of the impact at the firm-level whose results are broadly consistent with the notion that the impact of broadband expansion on imports operates through the alleviation of informational frictions. In Section 7, we introduce a firm-level model of input importing and derive the welfare formula allowing us to quantify the contribution of the import-channel as well as the total effect of broadband internet on consumer welfare.

2. Literature

This paper connects with several strands of the literature. First, it enriches the large literature regarding the interactions between trade and technical change. Second, it has implications for the origins of trade shocks whose effects have been extensively studied in recent years. Finally, it contributes to the literature on firm international sourcing choices.

While technology is often mentioned as a force amplifying globalization, empirical evidence on the matter appears scarce. For instance, in his rich account of the recent wave of globalization and its interplay with technology, Muendler (2017) mentions the possibility of technology-induced trade but does not cite empirical papers estimating the impact of ICTs on trade flows. Baldwin (2016) argues that globalization over the 1990s and 2000s was driven by information technology by lowering the cost of coordinating activities across borders. It boosted intra-industry trade between advanced economies and a handful of emerging

economies as multinational firms from the North moved labor-intensive tasks to the South. Steinwender (2018) provides empirical evidence on the role of the ancestor of modern ICT on trade during the 1800s by studying the impact of the telegraph on trade between the United Kingdom and the United States. A possible reason as to why modern data causal evidence on technology-induced trade is rather scarce might be related to the difficulty of measuring technology availability at the firm-level that varies over time while simultaneously observing trade patterns. Consequently, most studies documenting a positive impact of ICT on trade are either at the country (Portugal-Perez and Wilson, 2012; Clarke and Wallsten, 2006) or macro-regional level (Barbero and Rodriguez-Crespo, 2018).[3]

There are some recent exceptions that have documented the effect of technology using firm-level trade and outsourcing decisions. Kneller and Timmis (2016) find a positive effect of broadband internet on the export of services. Fort (2017) documents the firm-level relationship between investing in an electronic-integrated network and outsourcing, either domestic or international, of complex manufacturing tasks. In the context of developing African economies and relying on survey data, Hjort and Poulsen (2018) show that fast internet increases employment and that rising exports is one of the mechanisms explaining the upward shift in labor demand. Most closely related to our paper is Akerman et al. (2018) who estimate how broadband local availability affects the effect of distance on trade flows using a panel of bilateral trade flows between Norwegian cities and foreign countries, building on the identification strategy of Akerman et al. (2015). While their results display a significant interaction effect between fast internet and distance, they do not find that broadband increases significantly trade flows. Their coefficients are positive, although imprecisely estimated. We depart from these papers in several ways. First, we focus on the import side with the aim of quantifying the welfare consequences of the broadband-induced variation in imports. Second, we are able to document additional margins of adjustment of trade to broadband expansion, notably the "sub-extensive" margin of imports (i.e. the number of goods and origin-country imported by firms) that has been shown to be relevant in other settings (Gopinath and Neiman, 2014). Third, while Akerman et al. (2018) frame their empirical analysis in the gravity literature, aiming at explaining the distance puzzle, we adopt a more agnostic approach and aim primarily at documenting the extent to which technology-induced trade might have contributed to the recent rise in import penetration across advanced economies. Finally, we combine our estimates with a simple theoretical

[3]Several papers have recently documented the impact of trade on innovation, i.e. the reverse direction of causality compared to our focus. Aghion et al. (2018) find that expanding exporting opportunities increase the innovative activities of the most productive firms. On the import side, Bloom et al. (2016) find that greater Chinese import competition from China is associated with higher firm patenting in a panel of European firms while Autor et al. (2016a) find opposite results in the US case.

framework to quantify the impact of broadband internet on consumer welfare, relying on a sufficient statistic approach. We focus on welfare and do not model explicitly the impact of internet on the geography of trade.[4] A limitation of our paper with respect to Akerman et al. (2018) is that we do not observe firm-level *adoption* and therefore focus on local *availability* while they have access to information on broadband *adoption* from surveys. They can therefore estimate the effect of adoption on trade while we focus on the reduced-form. Note however that the reduced-form effect is relevant policy-wise as it is straightforward to increase local broadband availability through policy intervention while manipulating adoption involves more intricate interventions.[5]

Starting with Autor et al. (2013), a growing body of evidence, based on a local labor market approach[6], has documented the impact of trade shocks from emerging countries—in particular China and Eastern European countries, see Dauth et al. (2014)—on the local labor market outcomes (manufacturing employment, wages or employment rate etc.). Most studies analyze the effect of trade or technology in isolation, treating the factor left-out as a potential omitted variable to account for. By contrast, Autor et al. (2015) attempt to "untangle" the respective effect of trade and technology on several local labor market outcomes, notably the decline in manufacturing employment in the USA. Their analysis, however, considers sectoral aggregate trade flows as given. We instead make the point that the remarkable increase in trade flows over the period is in part driven by concomitant technical change. Moreover, we also find a large effect for imports sourced from Eastern European countries and China, suggesting that the recent "trade shocks" from the 2000s were, in part, driven by technological change. Our results echo the point raised by Fort et al. (2018) that in the presence of technology-induced trade, it is delicate to disentangle the effects of technology from those of trade.

A sizable reduced-form literature has documented that increased access to foreign inputs lead to favorable firm-level outcomes, notably productivity gains (Amiti and Konings, 2007; Topalova and Khandelwal, 2011). Most of the literature relies on trade liberalization episode to estimate the effect of trade in inputs on firm-level outcomes.[7] We contribute to this

[4]Note also that our sample size is considerably larger due to the sheer number of municipalities in France (about 35,000 cities versus 400 in Norway) thus allowing us to uncover effects that might be otherwise impossible to detect.

[5]See Andrews et al. (2018) for a cross-country empirical investigation of the numerous determinants of ICT adoption by firms and their complex interaction.

[6]This approach was pioneered by Topalova (2010) and Kovak (2013) who analyse the local impact of trade liberalization in India and Brazil respectively.

[7]In their recent assessment of the literature Shu and Steinwender (2019) review 20 papers, 15 of which adopt a reduced-form approach. Among these 15 papers, 14 rely on variation in foreign input access driven by trade policy reforms. The one exception is Juhász and Steinwender (2018) who estimate the effect of the

literature by using a new source of identifying variation (staggered roll-out of an ICT) and implementing a transparent event-study design that is allowed by the longitudinal nature of the data and the large number of local events. An additional contribution is to build on recent theoretical results in the literature on firm-level importing (Blaum et al., 2018) in order to translate our reduced-form estimates into an overall impact of broadband internet on consumer welfare and to isolate the contribution of the import channel.

As such, our findings also relate to the theoretical and structural literature studying the link between international sourcing, productivity and consumer welfare (Halpern et al., 2015). As broadband internet is found to increase the share of spending on foreign inputs as well as the variety of inputs, trade appears to be one of the mechanisms through which fast internet boosted productivity and ultimately consumer welfare. The paper documents the direct effect of broadband on firm-level outcomes, most importantly sales and value-added (subsection 6.3). In order to quantify the import channel through which fast internet boosts imports, we take our reduced-form estimates into a general model of firm-level importing. Rather than specifying the model fully and estimating its parameters (as in e.g. Sandoz, 2017; Antras et al., 2017), we rely instead on a sufficient statistics approach (Chetty, 2009) as first developed by Blaum et al. (2018) in the import literature.[8] Our estimates combined with the calibration of this model with French data implies that broadband internet allowed for a reduction of about 1.7% in the unit cost of production and that about 25% of that overall effect is related to the import channel.[9]

roll-out of the telegraph network on international trade. While the source of variation is similar to ours, the context (19th century), the level of variation (cross-country) and the outcome differ starkly from our set-up.

[8]These micro-results build on the seminal work of Arkolakis et al. (2012) that show that in a large class of trade models observed changes in the share of domestic consumption, together with the price elasticity of trade are sufficient statistics for evaluating welfare changes associated with a foreign shock.

[9]Perhaps more incidentally, our paper contributes to the literature on the overall impact of broadband internet (Bertschek et al., 2015) for a recent review. Previous works have notably documented the skill-bias of the productivity gains of broadband internet (Akerman et al., 2015; Ciapanna and Colonna, 2019). Few papers documented the local economic impact of broadband internet in France and no paper we are aware of look at firm-level outcomes. In a set of two papers, Houngbonon and Liang document a negative effect of broadband penetration, as measured by ratio of residential connections to the number of households, on the local Gini coefficient over the 2009-2013 period (Houngbonon and Liang, 2017) and a positive effect on service employment (Houngbonon and Liang, 2018). We instead focus on the period of expansion of broadband availability that is driven by the gradual upgrading of the copper wire infrastructure over the 1999 to 2007 period.

3. Data and context

3.1. *Context: the diffusion of Broadband Internet in France*

The ADSL technology. The ADSL (Asymmetric Digital Subscriber Line) is a data communication technology that enables fast data transmission over copper telephone lines (much faster than what a conventional voiceband modem could provide). In the ADSL technology, bandwidth and bit rate are said to be asymmetric, meaning greater towards the customer premises (downstream) than the reverse (upstream). Eligibility for ADSL depends on the distance between the final customer (e.g. home or office) and the Local Exchange (LEs), since the intensity and the quality of the analog signal decreases as it is routed over the copper lines.

Local Exchanges are the telephone exchanges owned by the incumbent operator France Télécom (later renamed Orange) into which subscribers' telephone lines end. As of 2008, there were about 17 000 LEs spread throughout the country. Initially dedicated to the telephone network, LEs are essential for Internet users who subscribe to ADSL. LEs aggregate local traffic and then direct it via the so-called backbone (i.e. higher levels of the network) towards the world wide web. A key feature of the ADSL technology is that one can supply high-speed Internet by upgrading the LE while relying on the existing (copper) local loop to connect the premises of the final customers. The upgrading involves the installation of an equipment inside the LE called a DSLAM (Digital subscriber line access multiplexer) that is required in order to recover the data transmitted via ADSL on the local copper loop and adapt it so it can be transmitted to the higher levels of the network (which are typically relying on optical fiber). The upgrading of local LEs is the key source of variation we will use in our empirical analysis (see section 3.2 for more details on the data).

The ADSL roll-out in France. The ADSL technology became popular during the 1990s, as many OECD countries were planning the expansion of services related to information and communications technology. In the early 2000s in France, the deployment of the technology beyond France's largest cities was slow. The causes for this staggered deployment are multiple. First, France Telecom (FT), the monopolistic telecom supplier at the time and still the main supplier today, was unsure as to whether it was going to be able to make the upgraded infrastructure available to new competitors with a positive markup or not. The uncertainty regarding the wholesale price FT was going to be able to charge made the firm reluctant to upgrade LEs beyond the largest cities (see Sénat, 2002, p.230). This uncertainty was

lifted after a series of decisions by the regulatory agency set the conditions of that wholesale market (Arcep, 2002).

Moreover, at the same time France Telecom had to invest massively in upgrading its LEs to ADSL, it went through an debt crisis which ended with what was essentially a government bailout in 2002. One can find anecdotal evidence of the impatience of the French government in accounts of Parliamentary debate (at the Senate) regarding the excessively slow expansion of broadband internet (Tregouet, 2001) and the difficult cooperation between the French government (the Ministry in charge of the Industry) and France Télécom.

Under the impulse of the government—which increased its stake in the firm during the 2002 bailout of the firm – France Telecom pledged in 2003 to cover 90% of the French (metropolitan) population by the end of 2005, i.e. all local exchanges (LEs) with more than 1000 lines, for a total investment of 600 M euros (750 M euros in 2018 prices) (Telecom, 2003).

Between 2004 and 2007, local governments (called départements[10]) started to play a role in subsidizing deployment and favoring competition among providers. Most relevant for broadband expansion is the creation of a contract between local governments the "Plan Département Innovant". It is a contract whereby France Telecom pledged to equip all LEs with more than 100 connections within a year in that département. The proclaimed target of the plan was to raise the French population coverage up to 96% by the end of 2005 and activate all the remaining LE by the end of 2006 (Telecom, 2003). About 50 departments signed the chart in 2004 even though many experienced delays. This commercial initiative was widely perceived as a way for FT to counter the direct involvement of départements in broadband provision that was made possible by a law voted in 2004.[11] We account for the role of départements in our empirical analysis by carrying a within-département analysis (we include département-year fixed effects throughout). FT claimed to have reached the full coverage of all relevant LEs in July 2007 (Le Gall, 2007), covering 98% of the French metropolitan population.

Overall, the account of the broadband expansion in France over the period suggests that it was gradual due to uncertainty regarding the capacity of France Telecom to undergo the investment until 2002. After 2002, with the strong impulse of the government, France Telecom started covering more secondary areas with a focus on the overall number of lines per LE with only limited attention paid to local economic potential. While accelerated, the coverage remained gradual due to operational limits on the part of FT and took about 2

[10] There 95 départements in mainland France.

[11] LOI n 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique

more years than anticipated in 2003.

Our main effects of interest are identified out of the gradual diffusion of the new technology in different LEs over space and time. The question of what were the criteria for deciding to "treat" one LE before another is of course central for our paper. It is the topic of a whole section immediately preceding the results (section 4.2), where we present evidence that the main determinant was the city-level population density, with no role for levels or trends in the trade patterns of the city.

Finally, it is worth noting that several sources show that the ADSL technology, while progressively replaced by other technologies—notably direct access to the optic fiber or FTTO (fiber to the office) –, is still the main way firms access the internet. In particular, based on a recent survey, it appears that 73 % of SMEs use ADSL technology only as of 2016 (Arcep, 2016). Moreover it constituted not only a massive improvement in terms of speed (from 56 to 512kbit/s for a transition from a classical to first generation ADSL connection) but also in terms of cost and time of connection.[12]

3.2. Data

In this paper, we combine three main sources of information: a unique city-level dataset on broadband internet availability, firm-level trade and employment, and firm-level balance sheet information. We provide a detailed description of our data in the following subsections. More details on the datasets used is provided in Appendix C.1.

Broadband internet data

The most novel aspect of our data is the (manually collected) date of upgrade to ADSL in mainland France for each Local Exchange (LE)'s.[13] The historical operator was compelled by law to make this data available to other operators as well as websites allowing consumers to gauge the quality of their line. The data was collected through one such website.[14] We additionally obtained data from the regulatory agency (ARCEP) regarding the geographical coverage of each LE. The data documents the area of each census block (IRIS) that is covered by a given LE. Each city in France is partitioned into census-blocks.

Combining both datasets, we construct a continuous measure of broadband access of city i

[12] FT was providing services at 128kbits at a much higher cost (Badré, 2007).

[13] Throughout the paper, broadband or ADSL refers to first generation ADSL that is associated with speed of 512 kbit/s.

[14] We were able to check with economists at Orange that our dates matched exactly their data for the years within the sample (1999-2007).

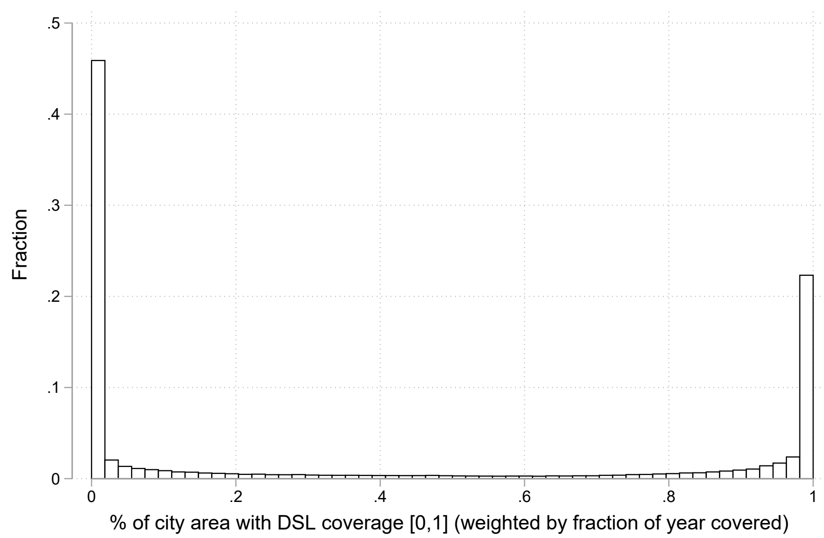
at year t . This measure, which we denote $\widetilde{Z}_{it}$, is a time-weighted percentage of area covered in city i. It is formally defined as:

$$\widetilde{Z}_{it} = \sum_{b \in i} \underbrace{\frac{\text{\# days with access in } b \text{ since Jan 1st of } t}{\text{\#days in year } t}}_{=D_{bt}} \times \underbrace{\frac{\text{area}_{bt}}{\sum_{b \in i} \text{area}_{bt}}}_{A_{bt}} \tag{4.1}$$

where $b \in i$ denotes the census tracks included in city i.

Discretizing the variable. We see that $\widetilde{Z}_{it}$ is in principle continuous between 0 and 1. $\widetilde{Z}_{it}$ will be equal to one if all of its areas have had access for the entire year. It will be equal to 1/2 if the entire city has had access to broadband over half the year t. While the continuous measure is useful, as it allows to gauge the state of broadband availability taking the area and time dimensions into account, we do not use it directly in our empirical estimations. Indeed, regressing trade outcomes on this measures would assume that the effect of an increase $\widetilde{Z}_{it}$ by, say 0.5, will be the same whether it stems from a coverage of the whole city over half the year or an increase in half the city area of the entire year.As we do not have strong theoretical reason to think that would be the case, but we still believe that the two sources of information are valuable, we discretize the treatment status by setting treatment status to 1 after the city experienced its highest increase in $\widetilde{Z}_{it}$. Formally, we define the year of treatment as $t_{i0} = \text{argmax}_t \Delta \widetilde{Z}_{it}$ and discretized treatment status as $Z_{it} = \mathbb{1}\{t \geq t_{i0}\}$.

Fig. 4.1. Distribution of $\widetilde{Z}_{it}$: 1999-2007



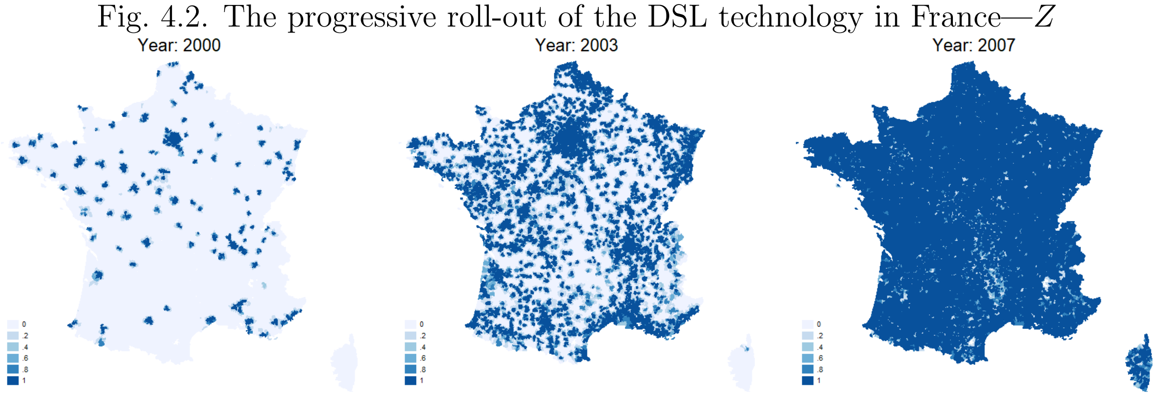
NOTES: This figure plots the distribution of the continuous measure of local broadband availability (variable $\widetilde{Z}$) as defined in Equation (4.1). We see that while the measure is continuous and contained between 0 and 1 but presents point of accumulation on 0 and 1.

As seen from Figure 4.1 in practice the empirical distribution of $\widetilde{Z}_{it}$ is heavily concentrated on 0 and 1 and accordingly Z_{it} and $\widetilde{Z}_{it}$ are strongly related ($\rho = 0.90$). Another way to assess the relationship between the continuous and discrete variables is to trace the evolution of $\widetilde{Z}_{it}$ around t_{i0}. Figure 4.15 plots the coefficients of a regression of $\widetilde{Z}_{it}$ on a set of dummies for each time with respect to t_{i0}. The results show a sharp increase between -1 and 0 from 0 to 0.4 and to 0.9 at $+1$. The coefficients are estimated with a very high degree of precision which reflects the fact that $\widetilde{Z}_{it}$ while continuous in theory has most of its support in 0 or 1.

Accordingly, the discretization of the variable does not result in much information loss. It is attractive as it allows us to use the more transparent event-study approach that will be described in Section 4.[15]

Geographical description of broadband expansion. Figure 4.2 shows the roll-out for all of France. Figure 4.21 in the Online Appendix focuses on a specific region (Occitanie) in order to provide a sense of the large variation in treatment status over time at very fine-grained geographical level. The dark areas represent a large degree of coverage (a high $\widetilde{Z}_{it}$). In 2000, those are confined to the few major cities of France, surrounded by a large majority of no-ADSL territories. By 2003, the treatment has largely spread to lower scale-municipalities, although large parts of France remain dependent on the old technology. The national territory is essentially "hole-free" in 2007, when almost all LEs have been dealt with. Our empirical approach will take account of the fact that all cities are eventually treated in our setup.

Fig. 4.2. The progressive roll-out of the DSL technology in France—$\widetilde{Z}$



NOTES: This figure presents the geographical distribution of the continuous measure of local broadband availability (variable $\widetilde{Z}$) as defined in Equation (4.1).

[15]The number of cities by cohorts can be found in Figure 4.16.

Trade and employment data

(i) Firm-level trade data. The data on firm-level trade is produced by the customs office, and compiles the exported values and quantities for each firm-destination-product combination over the period considered (1997-2007).

(ii) Firm-level employment and location data. The administrative dataset DADS (Déclarations Annuelles de Données Sociales) comes from firms' social security records. For the period 1997-2007 and for every establishment with at least one employee, it provides the number of workers, the overall wage bill and the city of location. This is an establishment-level dataset where a firm can have several establishments.

(iii) City-level covariates. Several covariates (population, share of college-educated workers, labor force) come from the 1999 census files aggregated at the city level and directly provided by the French statistical institute (INSEE) on its website. Additional covariates (number of households and overall tax income) come from tax authorities' files aggregated at the city level (Fichier communal de l'impôt sur le revenu).

Balance-sheet data

We use a firm-level balance sheet dataset from France. The firm-level accounting information is obtained from the BRN ("Bénéfices Réels Normaux") dataset which contains the balance sheet and income statement of all firms operating under the standard corporate income tax regime. The BRN dataset gathers information on about than 600,000 French firms every year. This dataset has been used in several trade-related papers dealing with French data (e.g. Eaton et al., 2011; Mayer et al., 2014). We use this dataset to compute firm-level performance measures (sales, value-added) as well as the relative importance of imports (imports-over-sales ratio). These measures will be particularly useful when linking our empirical results to the conceptual framework.

Construction of the estimating sample

We assemble those different data sources to construct a final city-level estimating sample. We start with the administrative dataset DADS to keep in our working sample only mono-city firms—i.e. firms whose all establishments are located in the same city. This is done in order to precisely identify the impact of broadband internet availability on trade outcomes (we want our treatment and outcomes to be at the same micro level). Indeed, while broadband internet expansion may occur at different moment for a same firm which owns several establishments in several cities, we only observe aggregate trade at the firm-level (the firm headquarter) and not at the establishment-level. Keeping only mono-establishment or

mono-city firms addresses this potential measurement issue. Mono-city firms accounted for 40% of the total value of imports as of 1999. Table 4.1 presents descriptive statistics at the city-level for our estimating sample. The average city hosts about 78 firms, 26% of which on average belong to the manufacturing sector. Imports are positive for about 50% of observations. The ratio of imports to overall sales is equal to 5% on average. Its distribution is highly skewed. Individual cities represent on average a small percentage of sales of overall sales (0.52%). This provides indirect support to the notion that broadband expansion of any given city is unlikely on average to have a major competitive impact on firms located in other cities.

Table 4.1: Descriptive statistics at the city-level: 1997-2007

	Mean	p50	p75	p95
Nb. firm (by city)	78	19	47	239
Sh. of mfg. firms	0.26	0.17	0.42	0.82
Import share (wrt. sales)	0.05	0.00	0.05	0.23
Share of domestic inputs	0.94	0.99	1.00	1.00
Foreign over domestic share	0.13	0.00	0.07	0.44
Sales as % of departement sales	0.51%	0.07%	0.29%	2.03%

NOTES : This table presents descriptive statistics at the city-level on our estimating sample for 18,332 cities over the 1997-2007 period (201,580 obs.). The cities in our estimation sample contain a total of 41,537 blocks. The share of importing cities is the mean of a dummy equal to one if at least one firm was an importer on year t in city i, and zero otherwise. The import share is the city-level ratio of the value of imports over sales

We compute the value of imports, the quantity of imports, the number of importing flows and the number of different products imported for mono-city firms aggregated at the city-level and for several geographic areas and products categories (as defined by the BEC). We merge this trade database with broadband internet data. We then aggregate firm-level balance sheet data from the BRN at the city-level and merge this additional information with our main database—we restrict the BRN sample to single-city firms present in the DADS. Finally, we use firm-level employment and location data, data on population, share of college educated workers, labor force from the 1999 census files as well as number of fiscal households and overall fiscal income from fiscal files aggregated at the city level to construct a large fiscal and demographic control database that we merge with our trade working sample. We keep cities that do not undergo a change in geography over the period.

4. Empirical approach

4.1. Baseline specification

Because all spatial units are ultimately treated, i.e. connected to broadband internet over the period we consider, we use the timing of treatment (staggered over time) in order to identify the main effects of interest. We estimate a dynamic specification where we allow the effect on a city i / year t, to vary with time-from-treatment. The year of treatment for each city is denoted t_{i0}. We index time-to-treatment with d (negative before treatment, positive after). Our sample covers the years 1997 to 2007, and we restrict the set of observations where $d \in \{-6, -5, ..., +4, +5\}$. The main estimating equation is as follows:

$$Y_{it} = \sum_{\substack{d=-5 \\ d \neq -1}}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0i}\} + \mathbf{x}'_{it}\boldsymbol{\delta} + \alpha_i + \psi_{r(i),t} + \varepsilon_{it}, \tag{4.2}$$

where α_i and $\psi_{r(i),t}$ are fixed effects for the city and for the département (of the city)-year, and $\mathbf{x}'_{it}$ is a vector of time-dependent city-level covariates. We drop two indicator variables for $d = -6$ and $d = -1$. That restriction is necessary to avoid multi-collinearity and to identify the fully-dynamic underlying data generating process in the staggered design (Borusyak and Jaravel, 2017; Gross et al., 2018). To ensure that this restriction is not influential in our results, we display results with alternative normalizations in the robustness section.[16]

The specification presented in equation (4.2) includes leads and lags. The inclusion of leads allows us to assess the presence of pre-trends. We also estimate a simpler "semi-dynamic" specification where only the lags of the treatment are included, as presented in equation (4.3):

$$Y_{it} = \sum_{d=0}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0i}\} + \mathbf{x}'_{it}\boldsymbol{\delta} + \alpha_i + \psi_{r(i),t} + \varepsilon_{it}. \tag{4.3}$$

The event-study coefficients $\widehat{\beta}_d$ estimated from equation (4.2) can be interpreted causally under the identifying assumption that, conditional on receiving broadband over the period considered and conditional on city fixed-effects, the *timing* of broadband roll-out is unrelated to the outcome. The presence of systematic local factors that would drive both broadband and trade would be cause for concern. This potential issue is investigated by assessing the

[16] We assess the sensitivity of our results to binning the coefficients associated with period -6 to -4 together, meaning to constraint them to be equal as in Schmidheiny and Siegloch (2019). Finally, note that the semi-dynamic specifications are not subject to this under-identification problem.

sensitivity of the coefficients to the inclusion of a large set of controls and fixed effects meant to account for city and well as local labor market shocks.

4.2. Validation of the research design

Explaining broadband expansion. Our identification strategy hinges on the assumption that the coverage of cities was mostly determined by city population density—which is mostly fixed over time—and did not take into account underlying trends in trading / importing activities. As a result, conditional on city and year fixed-effects, we consider the variation in broadband access to be as good as random. In order to assess the validity of this assumption, we explore the extent to which broadband coverage over time can be explained by different types of lagged city-level covariates. We group those covariates into several groups:

1. **Density**: population in 1999 per square km (log), interacted with a full set of year dummy variables.
2. **Industry dynamics**: shares of employment in 10 economic sectors at $t-1$ as well as changes in shares between $t-1$ and $t-2$.
3. **Trade**: asinh(number of transactions), asinh(value of imports) in $t-1$, changes in these two variables between $t-1$ and $t-2$.[17]

We estimate the following specification:

$$\tilde{Z}_{it} = \mathbf{dens}'_{it}\boldsymbol{\rho}_1 + \mathbf{indyn}'_{it}\boldsymbol{\rho}_2 + \mathbf{trade}'_{it}\boldsymbol{\rho}_3 + \mathrm{FE}_i + \mathrm{FE}_{r(i),t} + \varepsilon_{it}, \tag{4.4}$$

where $\tilde{Z}_{it}$ is the time-weighed share of city i that is covered by broadband internet as described in Equation (4.1). As we are mostly interested in the explanatory power of these different groups of observable variables, we only report the R-square of each regression. Individual coefficients can be found in the appendix.

Regression results. We start by regressing broadband internet coverage on all three sets of observable covariates without including any time or city fixed-effect. As indicated in Column (1) of Table 4.2, we obtain a R-square of 56% indicating that these variables capture a substantial share of the variation in treatment status. Column (2) presents the R-square of a two-way fixed-effect model including city and département × year fixed effects. This

[17] The inverse hyperbolic sine function (asinh) is close to the logarithm function and allows us to include city with 0 trade in the analysis. Amiti et al. (2019) is a recent example of a trade paper also using that transformation.

model absorbs 78.6% of the variance in treatment intensity. Column (3) presents the same model to which we add the 1999 measure of density interacted with year dummies. The fit of the model increases by 2.6 pp and the null that all the coefficients of the density variables are equal 0 are firmly rejected with a F-stat around 222. Interestingly, the set of industry dynamics or trade variables barely increases the fit of the model (columns 4 and 5) and the F-statistic of the joint test that all industry (column 4) or trade variables (column 5) are equal to zero are fairly low, indicating that, conditional on city and province-year fixed effect, they are roughly unrelated to the timing of internet coverage. Column (6) shows that the F statistics associated with the null hypothesis that all the industry and trade variables are null is small (2.5) and two orders of magnitudes smaller than that associated with population density (224). We consider the low predictive power of observable variables as supporting our identification strategy, since a large share of the variation in timing of the broadband expansion seems to be idiosyncratic in nature.[18]

Table 4.2: Explaining city broadband coverage: panel analysis

	(1) Covariates	(2) Twoway FE	(3) (2)+density	(4) (2)+indus.	(5) (2)+trade.	(6) (2)+ all cova.
R^2	0.557	0.786	0.812	0.787	0.786	0.812
Trade and/or industry: F-stat	59.56			2.57	4.78	2.5
Density: F-stat	20076.81		221.55			223.78

NOTES : This table presents the R-square of panel regressions following equation (4.4). Twoway FE (Column 2) refers to a twoway fixed-effect model with city fixed effect and département × year FEs. Density (Column 3) includes 1999 population density at the city level defined as #. of inhabitant divided by city area interacted with year indicators. Industrial structure controls (Column 4) include the lagged share and their changes of sectoral shares (nine sectors). Trade controls include the lagged log of trade value and number of flows (levels and changes). Column (1) includes all of the controls without fixed effects. Individual coefficients are reported in Table 4.12 of the Appendix.

Our baseline specification will allow for differential (linear and quadratic) trends in outcomes based on initial density. We will further assess the sensitivity of coefficients to the inclusion of observables and show that results are little affected by their inclusion. Naturally, this test does not imply that unobservables are not biasing our estimated coefficients. However for this to be the case, these unobservables should be time-varying, correlated with the timing of broadband expansion and yet uncorrelated to the rich set of observable variables included whose inclusion we show does not affect our estimated coefficients. While this remains a possibility we cannot completely rule out, given the battery of robustness tests we provide, we view it as very unlikely.

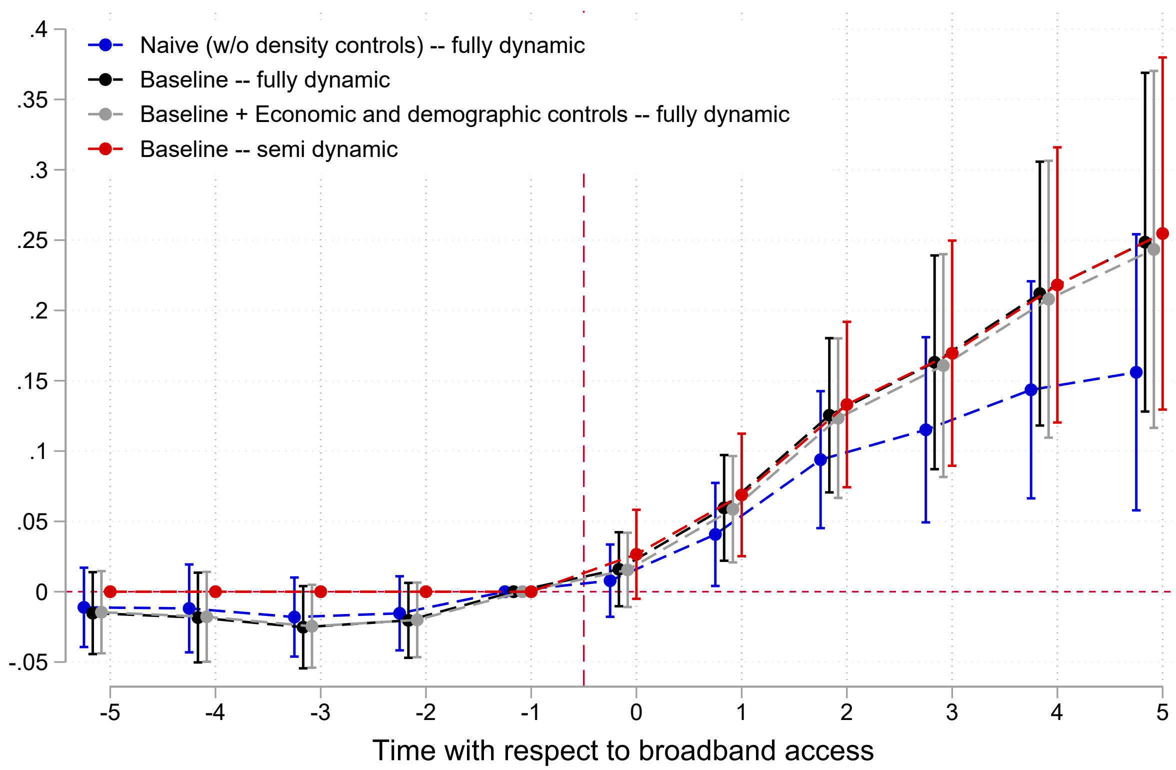
[18]Full regression results are displayed in Table 4.12. Trade and industry lagged level and dynamics are almost all insignificant once fixed-effects are introduced (column 6 of 4.12).

5. Baseline Results

The main variable we consider is the value of (in logs) imports of goods by firms located in a given city. We consider the overall value of all types of goods. We show the results for different specifications and assess the robustness of the results. We then turn to different margins of trade (extensive and sub-extensive margins).

5.1. Value of imports

Fig. 4.3. Main specification: Log of the value of imports



Notes: This figure plots estimates for specification in equation (4.2— fully dynamic) and (4.3—semi dynamic). The sample include all cities with a positive trade flow (import). The baseline specification includes 1990 population density at the city level interacted with quadratic and linear trends. Other controls include: (i) sectoral shares in 1997 (nine sectors) interacted with linear trend, (iii) the (log of) number of fiscal households and the average fiscal income in 1997 interacted with linear trend and (iv) 1990 education level interacted with linear trend (share of dropouts). 95 % confidence interval are presented. Standard errors clustered at the province (département) level. Full estimation results are reported in Table 4.6.

Figure 4.3 displays the main results of our paper, plotting estimated coefficients from Equation (4.2) (full results are presented in Appendix Table 4.6). The blue dashed line report results from a "naive" specification omitting controls except for the city and départementyear fixed-effects. Estimates exhibit a flat trend before the event (i.e. the normalizing measure of time since access $d = -1$) and a break in the trend after that. The coefficient

for $d = 5$ in that specification is 0.156 suggesting that the expansion of access to broadband internet increased the overall value of city-level import by about 16%, 5 years after the period of largest expansion (see column 2 of Table 4.6).

Given the relevance of density in the decision-making process leading to broadband expansion, our baseline specification adds linear and quadratic population density trends to the regression. The black line shows our estimates. Here again, we see no sign of a pre-trend prior to broadband expansion contrasting with a steady growth afterwards. The estimated effect after five years is substantially *larger* than in the naive case, with a coefficient of 0.249.

The coefficients displayed with a light gray line come from a third regression where controls for dynamics in sectoral composition, population and income (levels) as well as initial educational level (share of college-educated) interacted with linear trend. Results are remarkably stable to this inclusion and are very close to the specification including only density controls. The estimated lead coefficients appear equally supportive of the absence of pre-trends. Finally, the last set of coefficients plotted in red represent a semi-dynamic version of the baseline specification. The regression should in theory more efficiently estimated –as the number of parameters to be estimated is lower–, however the standard errors turn out to be very close to the fully dynamic specification in practice.

We see that the estimated effects (independently of the chosen specification) are growing with time since broadband internet expansion in a roughly linear fashion. The effects grow in magnitude as time passes which suggests that they are structural and might affect the path of the local economy and not simply its level. Given our reduced-form approach, this pattern might reflect the combination of two effects: the time lag between local availability of broadband and its adoption by firms as well as an internal lag between firm-level adoption and changes in importing behavior. In the absence of data on local adoption by firms, it is not possible to disentangle these two sets of effects. Our estimated coefficients reflect the products of both effects which might explain this pattern of linearly growing effects. Note that the time period included in our analysis ends in 2007. This choice is dictated by the advent of the financial crisis which has a major negative impact on global trade (Baldwin, 2009) and would add noise to the estimation.[19] Unfortunately it limits our ability to investigate the longer run effect of broadband expansion and whether it levels-off after a given period of exposure.[20]

[19]It is standard for empirical studies focusing on international trade to exclude the great recession from their time window, see for instance Autor et al. (2013)'s seminal study of the China shock in the US.

[20]Note however that documenting an effect at a 5-year horizon in a event-study setting is rather on the upper bound of what is found in the literature. For instance, the seminal study by (Autor, 2003) bins effects

Magnitude of the effect. In order to give insights about the quantitative implications of our findings, we construct counterfactual aggregate imports absent broadband expansion. The counterfactual outcome is measured as the actual outcome minus the predicted effect of broadband availability on the outcome, taking into account the dynamics of the effects as captured by our semi-dynamic specification.

More specifically, we compute an average effect of broadband internet expansion for each year effect as $\bar{b}_t = \sum_{t_0=1999}^{2007} w^y_{t_0,t'} \widehat{\beta}_{t-t_0}$ where $w^y_{t_0,t'}$ represents the national trade share for outcome y measured in t' for firms located in cities where broadband became available at year t_0. Let us denote the vector of total French imports over time as y_t. We postulate that the observed trade flow is given by a baseline level $y_t(0)$ that would have occurred in the absence of broadband diffusion multiplied by the predicted effect: $y_t = \exp(\bar{b}_t)y_t(0)$. We obtain the counterfactual series by inverting this relationship: $y_t(0) = \exp(-\bar{b}_t)y_t$.

We present two sets of counterfactual time-series in imports which corresponds to two different sets of weights w. The first is obtained by computing the weight for our estimating sample, which contains single-city firm only. The weights therefore do not sum to 1 for any given year but instead sum to the share of single-city firms in national imports.

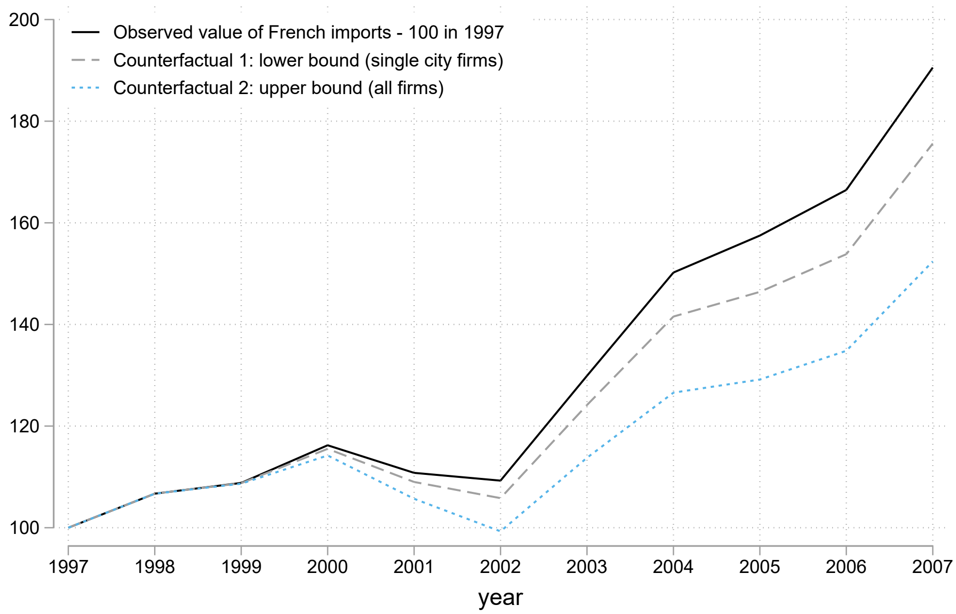
The second approach normalizes the shares just mentioned so that they sum to 1.[21] Applying the first set of weights implicitly assumes that multi-city firms were not affected by broadband expansion in their importing behavior, while applying the second set is equivalent to assuming that they reacted in the same way as single-city firms. Therefore we see the first counterfactual as a lower bound while the second is more likely to be an upper bound. Indeed, to the extent that multi-city firms are larger and might be able to invest in technology and commercial networks that decrease their reliance on broadband technology to engage in international trade, the impact of broadband internet on their importing behavior is probably lower than the same effect for smaller, single-city firms.

The results are displayed in Figure 4.4. The lower bound result suggests that the increase in the real value of French imports, which was 91% over the 1997-2007 period, would have been 15 p.p. lower without broadband expansion. The upper bound results suggests they would have been 37 p.p. lower.

4 years or more after the treatment.

[21] Formally, we define this second set of weights as : $\widetilde{w}_{t_0,t} = (\sum_{t_0} w^y_{t_0,t})^{-1} \times w_{t_0,t}$.

Fig. 4.4. Counterfactual aggregate trends in overall import



NOTES: The actual trade flows (black line) is the value of imports of goods in France in 2000 dollar normalized to 100 in 1997. Counterfactual 1 is obtained by subtracting the predicted effect of broadband internet where the average predicted effect is computed by using a weighted average of the estimated β_d, for $d \geq 0$ where the weights correspond to the share in national imports of each cohort of single-city firm (i.e. all single-city firms for which broadband expansion occurs the same year) measured the year of broadband expansion. The weights therefore do not sum to one and reflect the empirical importance of the estimating sample. Counterfactual 2 is obtained by doing the same calculation, but normalizing the previous weights so that they sum to one, thereby extrapolating the estimated effects outside of estimating sample onto all potential importers. See **Magnitude of the effect:** for more details.

5.2. Further Robustness checks

Local labor market dynamics. We saw above that including controls on sectoral employment dynamics leaves the results virtually unaffected. It could be however that our proxies do not fully capture local labor market shocks. In order to gauge the sensitivity of our results to the type of controls included, we control non-parametrically for any development at the local labor market level by including *commuting-zone* $\times$ year fixed effects. Commuting zone are defined based on a criterion of self-contained commuting and are the usual unit to study local labor markets in France. The estimated effects are displayed in Column 5 of Table 4.6 and are extremely close to the previous ones.

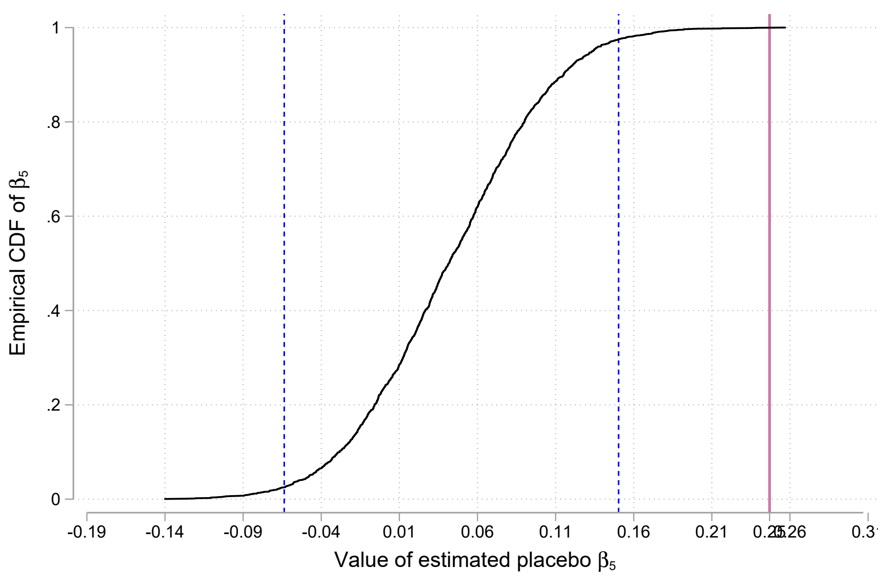
Alternative normalization. In the main analysis we normalize β_{-6} and β_{-1} to 0. Here, we experiment with a different normalization: instead of assuming $\beta_{-6} = 0$, we bin the coefficients associated with $d \leq -4$ together, that is we assume $\beta_{-6} = \beta_{-5} = \beta_{-4}$. Results are

displayed in Table 4.11. Again, focusing on column 3 (density controls only), the estimated effect is very close to the one obtained previously (0.30 versus 0.25). Overall, results do not appear sensitive to the choice of normalization.

Placebo inference. Bertrand et al. (2004) show that serial correlation can bias inference in difference-in-differences studies leading to serious over-rejection of the null hypothesis. We take this issue into account in our main analysis by constructing standard error that are clustered at the level of the département. These standard errors should lead to unbiased inference even in the presence of serial correlation within cities as well as cross-sectional dependence in the error term across cities within département.

In order to validate the inference provided by our clustered standard errors (which are only valid asymptotically) we implement Chetty et al. (2009)'s non-parametric permutation test of $\beta_d = 0$ for $d = 1, 2, 3, 4, 5$. To do so, we randomly reallocate the date of broadband expansion across cities, within the same département, and proceed to estimate equation (4.2). We repeat this process 2000 times and build an empirical CDF for $\widehat{\beta}_d$ which we denote $\widehat{F}()$. If broadband expansion has a truly significant positive effect on the dependent variable, here log of import values, one would expect the estimated coefficient to be in the very upper tail of the estimated empirical CDF based on permutations.

Fig. 4.5. Distribution of Placebo Estimates: Log Imports, β_5



Notes: This figure plots the empirical cumulative distribution function of placebo estimated effects broadband on log imports, where date of broadband expansion is randomly reallocated across cities within the same département (unit of cluster). Draws are with replacement and may include the correct date of treatment. The CDF is constructed from 2000 estimates of β_5 using the specification in equation (4.2) without observable controls. The solid line (in red) corresponds to the actual estimate of the matching specification. It lies outside of the 95% confidence interval that is delineated by the dashed lines (in blue).

Denoting $\widehat{\beta}_5^M$ the point estimate obtained in Figure 4.3 based on the log value of imports, we get $1 - \widehat{F}(\widehat{\beta}_5^M) = .0005$. Results are presented in Figure 4.5. This p-value is smaller than the one using the t-statistics based on asymptotically-valid clustered standard errors (0.001) and confirm that the broadband internet led to an abnormally large increase in the value of imports.[22]

Fixed-effect model with a continuous treatment variable. While discretizing the continuous treatment measure leads to very little loss in information due the underlying distribution of this variable (see Figure 4.1), our empirical setting lends itself well to a panel regression based on the continuous treatment measure. Here, we check that a panel regression based on the continuous treatment measure leads to qualitatively comparable results. As expected, point estimates are smaller since they measure a shorter-run effect but they are positive and significant at conventional levels of confidence (see Table 4.10 in the appendix).[23]

Inclusion of multi-city firms. Our baseline analysis focuses on single city firms—that is firms whose plants are located within a single city. This focus helps minimizing measurement error in treatment status given that trade is defined at the firm level but broadband expansion is defined at the city-level. This focus leads us however to discard a large part of the value of imports—as single-city firms accounts for about 40% of total French imports as of 1999. We check the sensitivity of our results to including multi-city firms.

The main difficulty is that we do not observe trade at plant or establishment level but at the firm-level. Accordingly, it is not easy to allocate firm-level trade to locations for firms with plants across several cities. We adopt two complementary approaches. The first one attributes all of firm trade to its headquarters. It is essentially equivalent to assuming that a firm is connected once its headquarters are. The second allocate firm-level trade to different cities depending on the weight of each city in its overall employment.

Results are displayed in Figure 4.19 of the appendix. It shows broadly similar responses when including or not multi-city firms with point estimates around 0.20 instead of 0.25 when focusing on single city firms.[24]

[22] The same holds for β_4, β_3 and β_2. Unsurprisingly, the null that $\beta_1 = 0$ is not rejected with as much significance but still above 5% (see Figure 4.18 in the Online Appendix).

[23] We also experimented with an alternative long difference specification that compares the change in imports from 1999 to 2007 between cities belonging to different groups of cohorts. As expected, we find that growth in imports tend to be higher in cities belonging to cohorts that received broadband in earlier years. Results as well as explanatory details of the approach are presented in Figure 4.25.

[24] This is consistent with an attenuation due either to measurement error but also to the heterogeneity

Accounting for zero-flows. Our results so far do not include observations where the log of imports are not defined because imports are null. As such, there might be an extensive margin at the city-level that we are missing. A simple way to accommodate observations where city-level imports are null is resort to the hyperbolic inverse function (denoted asinh) of imports on a balanced sample of cities.[25] Ashin() is defined as : $\text{asinh}(z) = \ln(z + \sqrt{1 + z^2})$.[26] The results are displayed in Table 4.9 of the Appendix. Coefficients are larger (0.50 for $d = 5$ instead of 0.25 for the baseline specification), suggesting a role for an extensive margin at the city-level and supportive of the causal interpretation of the results. We will therefore investigate in more details the margins of adjustment underpinning the overall positive effect of broadband expansion on imports in a dedicated section.

5.3. Intensive, extensive and sub-extensive margins

We have at this stage established a strong effect of broadband internet on the value of imports at the city level. This pattern could be consistent with an increase in the number of importing flows, defined as a firm-origin country-product combination, an increase in the average value per flow or any combination of both.

Figure 4.6 shows the results on both outcomes for the baseline specification. It is clear from those results that the average value per flow is virtually unaffected (coefficients are small in magnitude and lack statistical significance) and that the effect on overall value is almost entirely driven by the increase in the number of flows.

The increase in the number of flows could in turn reflect either an increase in the number of importing firms or a rise in the number of importing flows per firm. We show that the first impact (extensive margin at the firm-level) is small in magnitude (about 6%) albeit significant (see Figure 4.7). This implies that the second effect, the "sub-extensive margin" (Gopinath and Neiman, 2014), dominates: it explains most of the causal impact of broadband internet on the value of imports.

Our findings therefore support the notion that broadband internet caused an increase in import value, mostly by leading firms that were already importing to import more goods and

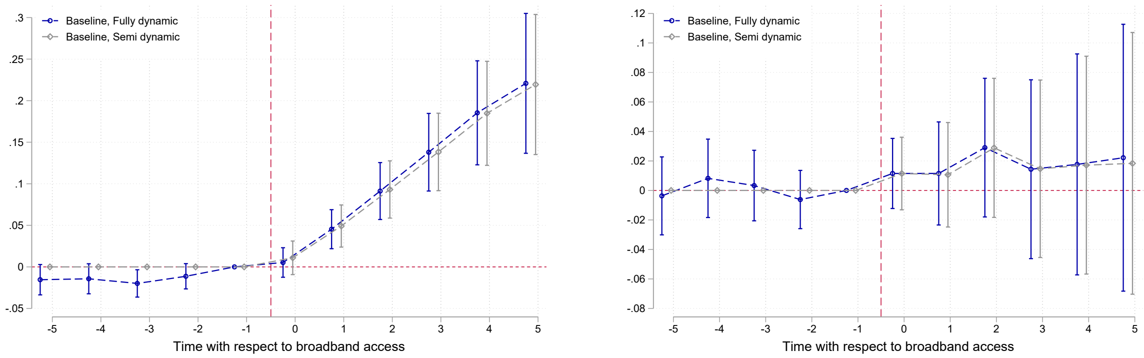
of the effect depending on size—see Table 4.3 in subsection 6.2 below. Note moreover that allocating trade proportionally to employment among establishment results in some differential pre-trends, plausibly reflecting the fact that headquarters matter most for firm importing decisions and that they tend to be affected by broadband expansion a few years prior to establishments in non-headquarter cities.

[25] For instance, Hjort and Poulsen (2018) use this transformation when studying how individual hours worked are affected by the arrival of fast Internet. Amiti et al. (2019) is a recent exemple using this transformation in an application aiming at quantifying the cost of recent US protectionist measures.

[26] It is such that for $z \geq 2$, $\text{asinh}(z) \approx \ln(z) + \ln(2)$ but with $\text{asinh}(0) = 0$

Fig. 4.6. Number of flows and average value per flow

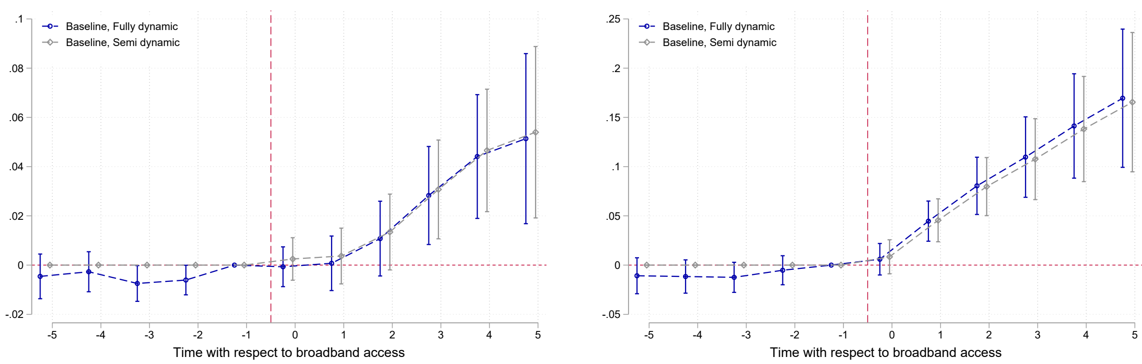
(a) Nb. of flows (b) Average value per flow



NOTES: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The number of flows (left panel) is defined as the number of the sum of all firm-origin-product combination for a given city in a given year. The average value of per flow (right panel) is defined as the overall value of imports by firms located in a given city divided by the number of flows as defined above. The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

Fig. 4.7. Extensive and sub-extensive margins

(a) Extensive margin (at the firm level) (b) Sub-extensive margins: flows per importer



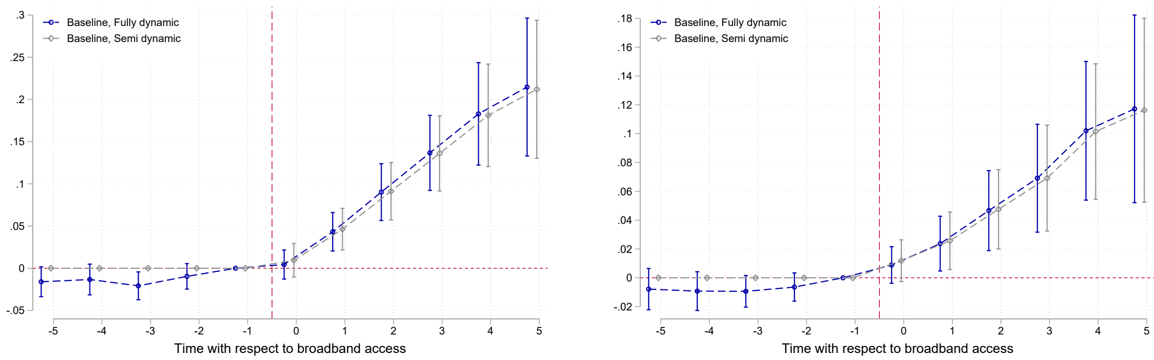
NOTES: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The extensive margin (left panel) is presented in The extensive margin (left panel) refers is measured by the log of the number of firms reporting positive imports at the city-level. The sub-extensive margin (right panel) is defined as the total number of flows, i.e. the sum of firm-country-product unique combinations reported within a given city, divided by the number of firms reporting positive imports as examined in the left panel. It proxies how diversified/large importers' sourcing set is on average. The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

from a wider array of origin countries while keeping the amount per flow roughly unaffected. This finding suggests that fixed costs play an important role in sourcing strategies (as in Antras et al., 2017) and also that broadband internet helps to decrease such costs. The value per flow being essentially unaffected, the increase in the amount spent on imports is roughly equal to the increase in the number of varieties imported—if one defines a variety as the unique combination of a product and a sourcing country as generally done in the

literature (Broda and Weinstein, 2006). Disentangling the rise in the number of products from the increase in the number of origin countries in the change in the total number of flows, Figure 4.8 shows that both margins appear relevant.

Fig. 4.8. Number of products (HS-6) and sourcing countries

(a) Number of products (HS-6) (b) Number of sourcing-countries



NOTES: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

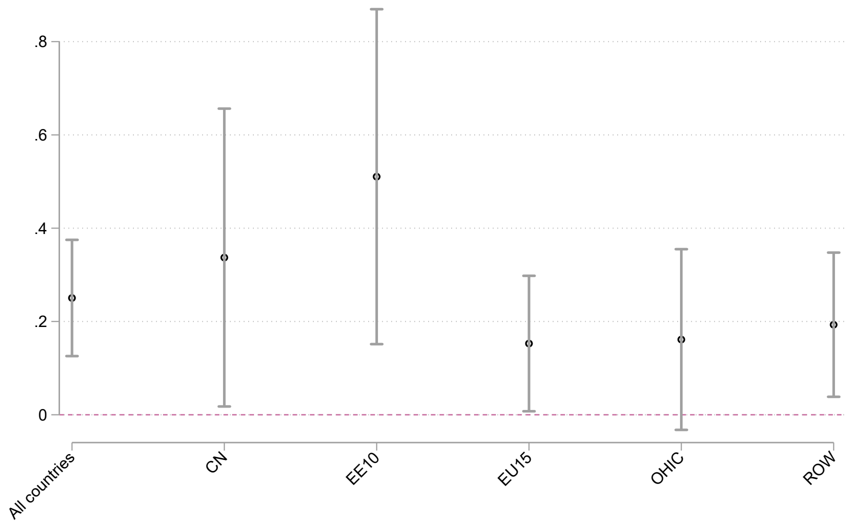
6. Heterogeneity, mechanisms, and margins of adjustment

In this section, we present further results regarding the heterogeneity of the effects by origin country and type of goods (subsection 6.1). We then present results on overall firm-level outcomes, namely sales and value-added (subsection 6.3). The latter results will be key inputs in our assessment of the impact of broadband on consumer welfare and the contribution of the import channel in the conceptual framework presented in Section 7.

6.1. Origin-country and type of goods

Origin-country. Figure 4.9 recalls the overall effect on the value of imports after five years and then shows the results for different groups of origin countries. Overall, we fail to detect statistically significant heterogeneity, although point estimates appear larger for imports sourced from China and from Eastern European countries and somewhat smaller for EU-15 as well as other high-income countries and the rest of the world.

Fig. 4.9. $\widehat{\beta}_5$ for different groups of origin-countries



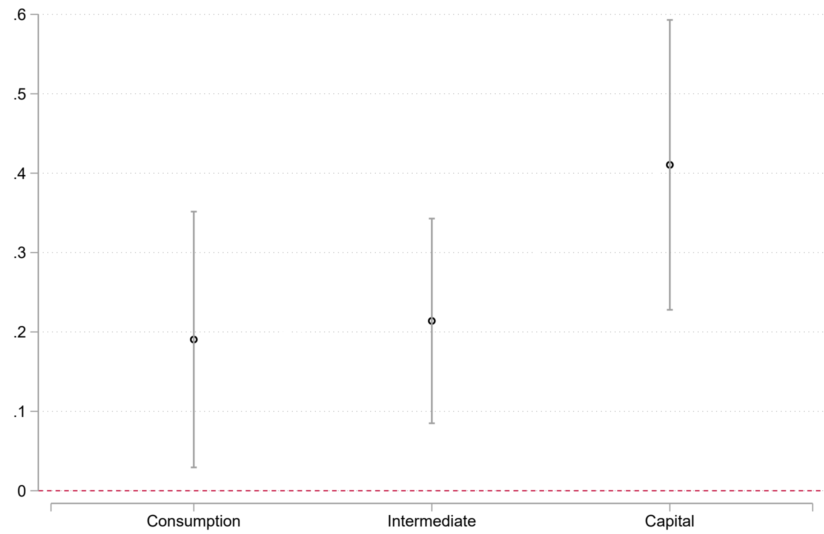
NOTES: This figure plots estimates $\widehat{\beta}_5$ from the specification in equation (4.3—semi dynamic) for different groups of countries. Density controls. The sample include all cities with a positive trade flow (import).. CN = China, EE10 = Eastern European countries that joined the EU in 2004. OHIC other high income countries. ROW = rest of the world.

Type of goods: capital, intermediary and consumption. Figure 4.10 shows the overall effect on the value of imports classified in three different types of goods based on the BEC classification: capital goods, consumption goods and intermediary inputs. Those results display some heterogeneity between types of goods: the value of imports of capital goods are strongly impacted while we find a smaller effect, albeit still positive and significant, for intermediary inputs and consumption goods.

We document further evidence of heterogeneity when considering combination of type of goods × origin country. The effects reported in Figure 4.20 of the appendix lack precision to reject equality of effect across countries within a product category. However, focusing on point estimates, we see that the results point to larger effect for China and Eastern European (EE10) countries than EU-15 countries regarding consumption goods. Regarding intermediate goods, the impact is higher for EE10. For EU15 countries, only capital goods are found to be significantly affected.

The literature on endogenous growth provides theoretical grounds for the role of foreign technology in enhancing domestic performance, notably through the import of capital goods (Eaton and Kortum, 2001). Our results imply that broadband internet might have boosted firm performance through that channel. To the extent that this channel is at work, one would expect firms sales and value-added to be boosted by broadband internet. Before we turn to such measures of firm performance, we investigate firm-level heterogeneity of the

Fig. 4.10. Value of imports by type of goods



NOTES: This figure plots estimates ($t \geq 0$) for specification in equation (4.3—semi dynamic) and for goods grouped. Density controls. The sample include all cities with a positive trade flow (import).

effect on imports.

6.2. *Analysis of heterogeneous effects at the firm-level*

Heterogeneity based on firms characteristics. In this section, we attempt to understand better the mechanism at play by documenting the heterogeneity by sector and other firm-level characteristics. While baseline regressions are based on city-level observations, results presented here rely on firm-level regressions. Our estimation sample includes all single-city firms with at least one positive trade flow (import) before Broadband Internet access. In order to summarize parsimoniously the heterogeneity in the effect we present here results based on a simple panel regression. The results are therefore most directly comparable to those presented in Table 4.10.[27]

Table 4.3 documents heterogeneity in the size of the effect by sector and firm's size and productivity. Column (1) present the average effect associated with the continuous treatment. The size of the effect is (+0.033) is comparable to what is city-level observations (+.049, see 4.10, column 4). Considering the effect by sector, it appears that most of the impact is driven by firms operating in the manufacturing (column 2) as well as retail sector (column 3). The

[27] Moreover, in order to confirm that going from city- to firm-level regressions does not affect the size of the coefficients, we provide event-study results based on firm-level regressions in Figure 4.24 of the online appendix, consistent with city-level estimates. Coefficients are somewhat smaller (+.15 versus +.25) but qualitatively similar.

other sectors show much smaller and insignificant effect (column 4). The remaining columns suggest that the effect is stronger for firms belonging to the bottom half of the distribution of size and productivity (as measured by value-added per worker). This could potentially reflect the fact that large and productive firms had already invested in their own IT internal capabilities and are less reliant on the state of local infrastructure. Another possibility is that these firms had already formed a network of foreign suppliers and were less reliant on broadband to overcome informational frictions impeding their matching with foreign suppliers. Note that, because larger and more productive firms tend to be more import-intensive (Bas and Strauss-Kahn, 2014, document this using French data), the documented pattern of heterogeneity implies that the ascent of broadband has contributed to allow smaller and less productive firms to import more and thus lower the correlation between size and productivity and imports intensity.

Table 4.3: Heterogeneity depending on sector and firm size and productivity

	All firms	Sectors			Productivity		Size	
		Manuf.	Retail	Other	Above median	Below –	Above median	Below –
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\tilde{Z}_{it}$	0.033***	0.038**	0.040**	0.003	0.021	0.039**	0.024	0.038**
	(0.011)	(0.015)	(0.017)	(0.035)	(0.019)	(0.019)	(0.015)	(0.018)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year x Dep. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Density controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	455485	169862	214055	68941	206585	165168.0	198504	173261

NOTES : This table presents results of a panel of firm-level regression where the explanatory variable of interest *Broadband Internet access* is the continuous measure of broadband access in city i and year t defined as a time-weighted percentage of area covered in city i as defined in equation (4.1). *Sample*: Firms included in the sample are all single-city firms with at least one positive trade flow (import) before Broadband Internet access. Robust standard errors clustered at the province (département)-level are presented in brackets. *Definition of variables*: productivity and size are defined as value added per worker andf employment the year prior to the broadband expansion. The median for both variable is constructed for the same year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See the figure 4.3's notes for more details on the controls.

Heterogeneity based on goods and trading partners. It is highly plausible that the reduction in the fixed cost of importing new products reflects lower informational frictions. We try to test this assertion by assessing whether the marginal effect of broadband coverage is stronger depending on the characteristics of the sourcing country as well as the good being sourced which are proxy for high initial return to reduction in informational frictions.

The impact of accessing broadband internet from the home country (France) on the amount of information gathered on a specific sourcing country is likely to be higher for sourcing country with high broadband coverage. Therefore, to the extent that the broadband-induced reduction in informational frictions between locations is higher when both locations are well covered by broadband, one would expect that the effect of broadband coverage on imports

is stronger for goods sourced from high coverage countries. This is what we find when comparing column (2) versus column (3) of Table 4.4.

Rauch (1999) and subsequent work (e.g. Nunn, 2007) argues that informational frictions are more important in determining the matching international buyers and sellers and shaping trade flows for differentiated products than for products traded on organized exchanges. Therefore to the extent that broadband internet stimulates imports through the alleviation of informational frictions, one would expect its marginal effect on average to be higher for differentiated than for homogeneous products. As reported in column (4) and (5) of Table 4.4, we find results broadly in line with this prediction.

Table 4.4: The role of information: Heterogeneity based on origin and products characteristics

	All Goods	BI coverage in sourcing countries		Differentiated Goods (Rauch)	
		High BI countries	Low BI countries	Differentiated	Not differentiated
	(1)	(2)	(3)	(4)	(5)
$\tilde{Z}_{it}$	0.033***	0.045***	0.032*	0.023**	0.009
	(0.011)	(0.013)	(0.019)	(0.012)	(0.032)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year x Dep. FE	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Density controls	Yes	Yes	Yes	Yes	Yes
Observations	455485	424425	248520	433544	123775

NOTES : This table presents results of a panel of firm-level regression where the explanatory variable of interest *Broadband Internet access* is the continuous measure of broadband access in city i and year t defined as a time-weighted percentage of area covered in city i as defined in equation (4.1). *Sample*: Firms included in the sample are single-city firms with at least one positive trade flow (import) before Broadband Internet access. Robust standard errors clustered at the province (département)-level are presented in brackets. *Definition of variables*: Country level broadband coverage is defined using the World Development Indicators database. A country is defined as high coverage if it has a coverage above the median as of 1999. The definition of differentiated goods is obtained from the classification proposed by Rauch (1999). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See the figure 4.3's notes for more details on the controls.

Overall, while inferring mechanism from observational data is necessarily delicate, we view the firm-level evidence gathered in this section as strongly consistent with and broadly supportive of the notion that BBI boosted imports through reduction in informational frictions.

6.3. Impact on firm performance and import-intensity

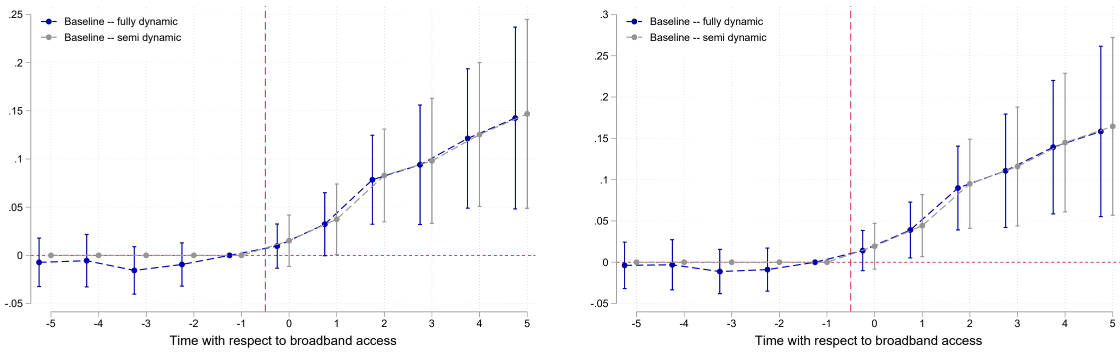
Import-intensity. We have seen that broadband internet increased the overall value of imports. In this section, we assess whether it led to an increase in *import-intensity* which normalize the value of imports by some measure of the size of economic activity undertaken by firms located in a given city.

We use two measures which leads to broadly similar results. The first variable measure imports over sales while the second measures what we call the "foreign share" which is the ratio of import to overall inputs. Results are displayed respectively in panel (a) and (b) of

Figure 4.11. We see in both cases a positive coefficient of around 0.15 after five years.[28]

Fig. 4.11. Measures of import intensity

(a) Imports-over-sales (log) (b) Foreign share of inputs (log)



Notes: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). Ratio at the city level are computed as the weighted average of firm level ratios for firms located in a given city, where weights are the denominators used the ratio—value of sales and inputs for the left and right panel respectively. The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

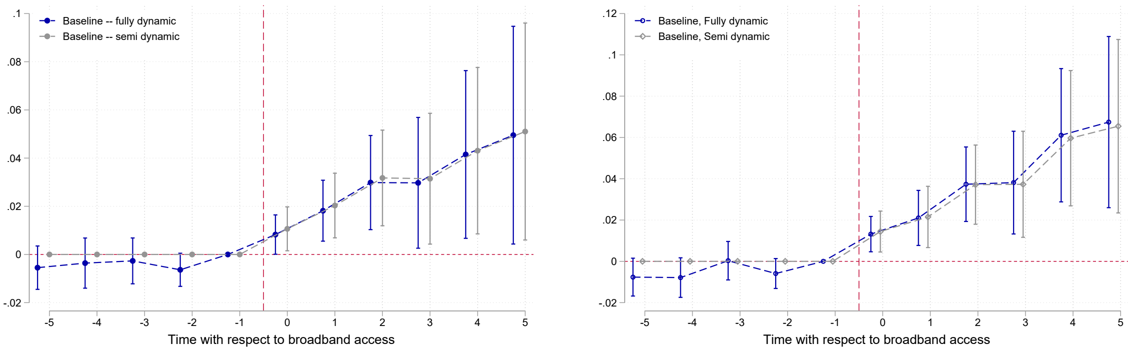
Firm performance 1: Value-added, sales and productivity. While trade is an important outcome in its own right, it is also interesting to see whether the increased imports of firms is associated with an expansion of overall activity as captured in their sales or value-added. The results, displayed in Figure 4.12, show increases by 5% in sales and 6% in value-added.

Beyond the scale of economic activity, access to a broader range of inputs have been documented to increase productivity (Amiti and Konings, 2007; Topalova and Khandelwal, 2011). As most of the evidence focuses on the manufacturing sector (e.g. Halpern et al., 2015; Blaum et al., 2018) we split the analysis of productivity (value added per worker here) between the manufacturing and the non-manufacturing sectors.

[28] The foreign share at the city-level is computed as the weighted average of firm-level foreign share of inputs where weights are firm inputs. That is the share is computed as follows: $sf_{it} = \sum_{j\in i} \frac{IC_{jt}}{\sum_{i\in j} IC_{jt}} \times \frac{M^n_{jt}}{IC_{jt}}$ where IC_{jt} is the value of overall intermediate consumption as reported in the balance sheet data and M^n_{jt} is the value of imports as reported in the customs data for firm j at time t. We assess the robustness of our results when excluding capital goods from the customs data in order to make the measurement from the two data sources more comparable. Indeed, measuring capital goods in the balance sheet is delicate because they are likely to be accounted for as investment rather than as intermediate consumption and there is no distinction between domestic and foreign goods in the data at hand. Excluding capital goods allows us to make sure that the imports as documented in the customs contain goods comparable to intermediate consumption as documented in the accounting data. Results excluding the capital goods are broadly similar with a point estimate of +0.147. See Figure 4.23 in the appendix.

Fig. 4.12. Sales and value-added (log)

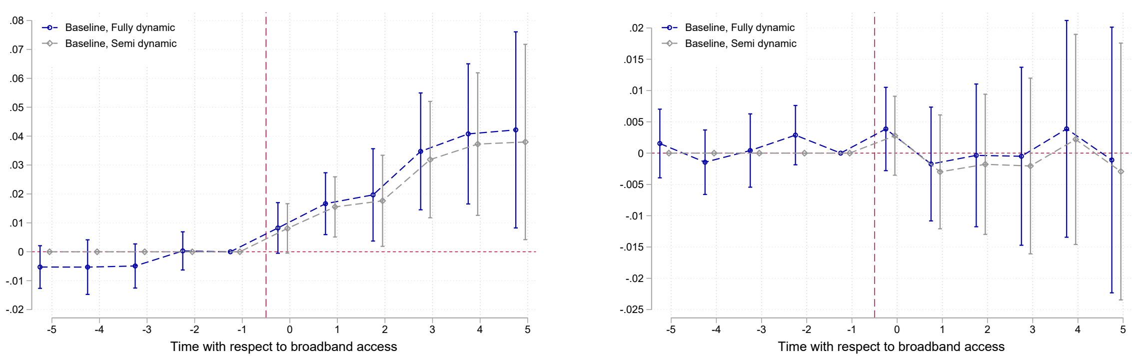
(a) Domestic Sales (b) Value-added



NOTES: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

Fig. 4.13. Value-added per worker (log)

(a) Manufacturing sector (b) Non manufacturing sector



NOTES: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

Results in Figure 4.13 shows a positive effect of broadband expansion on the productivity of the manufacturing sector, but essentially no impact outside of the manufacturing sector where value-added and employment grew in the same proportions. The positive impact on firms' scale with no effect on the apparent productivity of labor in the non-manufacturing sector is consistent with previous work for the UK provided by DeStefano et al. (2018) who find a positive average treatment effect of ICT instrumented by local broadband availability on firm's revenue and employment but not on different measures of productivity.[29]

[29]They do not show long run effect on productivity by sector however.

Firm performance 2: Exports. While our paper deliberately focuses on imports as it matters most directly for consumer welfare and is an understudied outcome relatively to exports, there are reasons to believe that exports might also react positively to the improvement of information infrastructure following broadband expansion. This is broadly what we find although results are less precise than in the case of imports.

Panel (a) of Figure 4.14 shows that export grew faster post broadband expansion, with a coefficient of about +0.20 after five years. In the light of the positive impact we obtained on overall value-added and sales, the fact that exports increases could be a mere reflection of overall gains in productivity with no specific effect on exports through reduction in trade frictions.

To investigate whether there is a specific impact on exports, that is a reduction in the cost of exports above and beyond unit cost reduction, we investigate the effect on the export over sales ratio. The results are displayed in panel (b) and point to a positive effect although this coefficient. Quantitatively, the coefficient is very similar to the one regarding the import-over-sales ratio for $d = 5$ (0.125 versus 0.15).

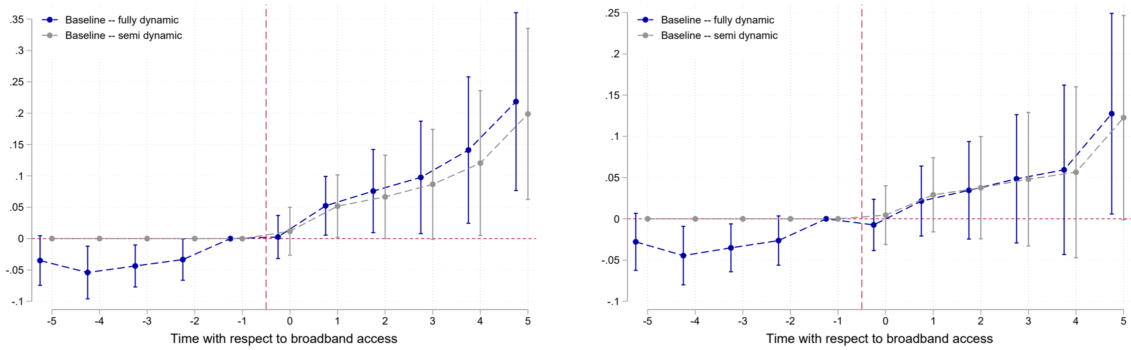
We see however that the coefficient is not very precisely estimated and is significant at the 5% level for $d = 5$ only. We notice moreover some differential pre-trends suggesting that a causal interpretation of the coefficients might be more difficult than in the case of importing which shows a clear break of trend around the expansion year (see Figure 4.3 and 4.11 above).

This could be lower precision of our estimates could reflect a higher degree of heterogeneity among exporters than importers in their reaction to the broadband expansion or that in our particular setting, the information frictions alleviated by the particular technology we consider are more relevant for importers than exporters. This is consistent with the fact highlighted in survey data from the early 2000s that firms use internet mostly to buy rather than to sell and that this is particularly true for small and medium enterprises (OECD, 2004).[30] In view of the more mixed results we obtain for exports, our modelling exercise in the next section will ignore the export channel in its basic version but will allow for it as an extension.

[30] See in particular Figure 3 and 4, page 15-16.

Fig. 4.14. Exports

(a) Exports (log) (b) Exports-over-sales (log)



Notes: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (export).

7. Conceptual framework for quantification

In this section, we set-up a simple but general model of firm-level imports in order to assess the welfare implications of our findings. We aim in particular to quantify the *overall* impact of broadband internet on consumer prices and the contribution of *import channel*, i.e., the increase in consumer surplus that can be attributed to how broadband increases import penetration.

The full derivations of the model are contained in the online appendix (section B). Here, we present the building blocks of the models, alongside the main formulas and intuitions allowing us to relate our empirical estimates to changes in consumer welfare. We also discuss the key assumptions of the model and how relaxing them would change the interpretation of our results.

7.1. Baseline model

Production side. The first building block relates to the production side. Firms combine labor and intermediate inputs using a constant return to scale Cobb-Douglas production function—with a labor share denoted as γ. Intermediate inputs are themselves a CES aggregate of domestic and foreign inputs—with an elasticity of substitution denoted as ε. Firms are heterogeneous in terms of TFP as well as in the intensity of domestic inputs use. Firms choose their bundle of inputs and their optimal sourcing strategy—i.e. the set of product-country combinations they source from—in order to minimize their unit cost of production.

In this environment, we allow broadband expansion to affect firms along several margins. Broadband can directly affect TFP, reduce the effective price of labor through skill-biased technical change (SBTC as documented in Akerman et al. (2015)) and lower the price of imports. The model remains agnostic regarding which type of trade costs (variable, fixed per destination or product, search friction etc.) is affected by the broadband internet expansion shock.

Under this set of assumptions, we can express the total change in unit cost induced by broadband expansion as the sum of the *import channel* and a *residual term* that includes all the other channels. Denoting $u_i(0)$ and $u_i(1)$ the unit cost for firms in city i without and with access to broadband internet respectively, we write:

$$\mathbb{E}[\ln(u_i(1)) - \ln(u_i(0))] = \text{Import channel} + \text{Other channels } (\Delta \text{ in TFP}) \tag{4.5}$$

Moreover, we can express the import channel—i.e. the unit cost effect of BI through changes in the import environment—solely as a function of (i) the effect of BI on the foreign share of inputs and (ii) the parameters of the production function (γ and ε). Indeed, the combination of cost minimization and price-taking behavior implies that the share of inputs sourced abroad is a sufficient statistic for the effect of trade in inputs on unit costs. Accordingly, the impact of broadband internet on the foreign share—whose estimates are presented in Figure 4.11—is a sufficient statistic for the part of the effect of broadband internet on the unit cost that is channeled through changes in importing behavior. The formula for the import channel can be expressed as:

$$\text{Import Price Channel} \equiv \mathbb{E}\left[\ln\left(\frac{u_i(1)}{u_i(0)}\bigg|_{\text{TFP fixed}}\right)\right] = \frac{1-\gamma}{\varepsilon-1}\mathbb{E}\left[\ln\left(\frac{1-sf_i(1)}{1-sf_i(0)}\right)\right], \tag{4.6}$$

where sf_i stands for city i firms' share of foreign inputs. Equation (4.6) states that the broadband-induced decline in unit cost through changes in the import environment is directly related to the decrease in the domestic share (i.e. $1 - sf_i$) with a coefficient depending on parameters of the production function. A given change in the domestic share will translate in larger variation in the unit cost for lower values of both the labor share γ and the elasticity of substitution between foreign and domestic inputs ε. At this point, it is worth emphasizing that the translation of BBI-induced changes in domestic changes into unit cost of production does not require any assumption regarding the market conduct of firms or the demand system for the final good.

Demand-side, type of competition and the firm location problem. To translate the decline of unit costs into consumer welfare we need to specify consumer preferences as well as how firms compete. We consider the case where a representative consumer has CES preferences over the varieties produced by each firm—with elasticity of substitution $\sigma > 1$. We further assume that firms operate under monopolistic competition.

The French economy is made up of a large set of cities which differ in terms of their local productivity and cost of importing. Trade costs between those cities are null. Firms' productivity depends on an idiosyncratic component which varies at the city $\times$ firm level and is distributed Fréchet—with shape parameter $\theta > \sigma - 1$. Firms accordingly pick the profit-maximizing location among French cities.

Monopolistic competition and CES preferences combined with constant return to scale ensures the full pass-through of variations in unit costs into prices. Under the additional assumption that the impact of broadband on the foreign share of inputs is homogeneous across firms, the model predicts that the percentage decline in the consumer price index is identical to the unit cost decline experienced by each firm (see equation 4.18 in the appendix).

This set of assumptions further allow us to back-out the overall impact of broadband on unit cost—i.e. the impact inclusive of all the channels including the import channel—from the effect of broadband on city-level sales, i.e. sales realized by firms located in a given city (independently of where the customers are located in France), whose estimates are presented in Figure 4.12, panel (b). More precisely, we can express the overall effect as:

$$\text{Total Price Channel} \equiv \mathbb{E}\left[\ln\left(\frac{u_i(1)}{u_i(0)}\right)\right] = -\frac{1}{\theta}\mathbb{E}\left[\ln\left(\frac{x_i(1)}{x_i(0)}\right)\right] \tag{4.7}$$

where $x_i(0)$ and $x_i(1)$ denote city i firms' sales without and with access to broadband technology respectively. The average effect of broadband on the local unit cost of production can be backed-out from the local average effect on sales, given an assumed value for θ—the Fréchet shape parameter that is inversely related to the variance of the idiosyncratic productivity component.[31]

[31] Intuitively, when θ is small, the variance of the idiosyncratic component is large, so that a local increase in the non-random part of productivity will not attract many firms. Inversely, when θ is large, firms are almost indifferent across cities so that small changes in local productivity will generate large-scale relocation of firms, and also boost sales by local firms. It might appear surprising that the elasticity of sales x to unit cost u at the city-level does not depend on the elasticity of substitution σ but only on θ. This is however a standard feature of models featuring discrete and continuous choices. In models following (Hanemann, 1984), the value function derived from the continuous choice problem (here the expected profit given a specific location) depends on an idiosyncratic term which then drives the discrete choice (here location choice). With Fréchet or Gumbel-distributed heterogeneity, the parameter driving *directly* the continuous choice (σ) cancels out—see equation 4.14 in the appendix.

Main results. Under the assumptions detailed above, we combine our event-study estimates on sales ($\widehat{\beta}_5^{\text{sales}}$) and on the foreign share ($\widehat{\beta}_5^{\text{sf}}$) with sample statistics on the average value of the pre-broadband foreign share of imports ($m \equiv \overline{\frac{sf(0)}{1-sf(0)}}$) and with calibrated values of three parameters of the model $(\theta, \varepsilon, \gamma)$ to quantify the change in consumer prices allowed for by broadband expansion.[32]

Our estimates for the total impact of broadband on consumer prices and the import channel write as follows:

$$\underbrace{-\frac{1}{-\theta}\widehat{\beta}_5^{\text{sales}}}_{\text{total}} = \underbrace{\frac{1-\gamma}{\varepsilon-1}\widehat{\beta}_5^{\text{sf}} \times m}_{\text{import}} + \text{residual effect.} \tag{4.8}$$

Given the calibrated values and our event-study point estimates we obtain:

Total Price Channel $= -1.85\%$, and **Import Price Channel** $= -0.75\%$.

Given our calibration, we find that the import channel led to 0.75% increase in consumer welfare through its effect on the price index. As a point of comparison, Berlingieri et al. (2018) estimate in a recent paper that trade agreements implemented by the EU led to an overall decline of 0.24% in the consumer price index over a 20 year period (1993-2013).[33] Therefore, while the import channel of broadband internet access might at first seem moderate, it is large compared to other policy experiments and accounts for about 40% of the consumer gains from broadband internet.[34]

[32] Value for the parameters θ, σ and ε are taken from the literature. Parameter θ is obtained from the most comparable estimate of Fajgelbaum et al. (2019), which also model it as a firm-location productivity dispersion parameter. For the value of ε, we rely on Blaum et al. (2018)—whose estimation is also carried out on French data for a comparable period (2000-2006). Our preferred estimate of σ is the one using firm-level intensive margin trade elasticity by Berthou and Fontagné (2016). The parameter $1-\gamma$ (intermediate good elasticity of output) is backed-out from the average ratio of intermediate spending over sales after correcting for mark-up. A full summary of the sources is provided in Table 4.7 of the appendix.

[33] Note that we use consumer price and consumer welfare interchangeably here. This is an approximation as some of the fixed costs firms need to spend to engage in importing might reduce firm profits. If consumers own the firms, their welfare will be negatively affected as their nominal income from firm ownership will decrease. Therefore reduction in consumer price will be an upper bound for gain in consumer welfare. We consider however this effect to be likely negligible. The model presented here is partial equilibrium in the sense that we ignore the link between firm profits and consumer income. Moreover, we take the price of domestic intermediates as fixed and we do not impose a balanced trade condition. We show in Online appendix E.1 that allowing the price of intermediates to be endogenous does not subtantially impact the welfare analysis.

[34] When ignoring capital goods, the import channel is somewhat lower with an effect of -0.61% on the price level, i.e. around 30% of the total effect. This lower effect reflects both a slightly lower estimate for $\widehat{\beta}_5^{\text{sf}}$ (0.148 instead of 0.158) and a lower m (0.101 instead of 0.13). More details regarding the calibration are provided in appendix Section B.5 and in particular Table 4.7.

7.2. Extension: exports and labor market clearing wage

Discussion of main assumptions. The quantification described so far is in partial equilibrium in that it takes wages as exogenous and consider that the broadband shock is too small to change their equilibrium value. We further ignored the impact of broadband on exports. The two are linked: export performance can matter for workers' welfare to the extent that its raises wages. Accordingly, here we jointly relax two assumptions:[35] firms are now allowed to export more in response to broadband and wages adjust at the national level in order for the overall labor market to clear.

The export wage channel. Allowing for endogenous wages and exports results in a new channel of influence of broadband on welfare. We call this impact the *export wage channel*. This channel will not leave the channels we saw in the previous subsection unaffected. In particular, it will change the effect of broadband on consumer prices to the extent that aggregate wage increases are passed-through into prices.

In section B.6 of the online appendix, we introduce exports in a very simple way: firms choosing to locate in city i export to foreign market k and face an iceberg cost τ_{ik}. Broadband reduces τ_{ik} and consequently increases the local *market potential* (i.e. the demand to be expected when locating in i). Labor is perfectly mobile across cities and cities are small relative to the overall economy so that the local impact of broadband has no direct impact on local wages. However, a full nation-wide broadband expansion raises the demand for labor and therefore pushes the market-clearing wage up. The nationwide labor market clearing condition gives rise to a wage equation, in which wages are affected by two terms: (i) positively by market potential (itself increased through broadband-induced reduction in τ_{ik}), and (ii) negatively by non-wage determinants of the unit cost (which are impacted by broadband through productivity gains or by the import-channel).[36]

Note that a maintained assumption is that null trade costs within France. This assumption allows us to separately identify unit cost reduction and increase in (export) market potential. We retrieve the former from the increase in *domestic sales*, i.e. output of a city that is sold to French consumers. Moreover, we show that we can retrieve the increase in (export) market potential in city i from the increase in the *share of export in total sales*. Our event study provides an estimate for such an effect—see Figure 4.14.

Based on this additional structure, we propose two decomposition of the aggregate welfare

[35] All details of the model can be found in section B.6 of the appendix.

[36] See equation 4.42 in the appendix B.6 for a formal breakdown of the wage effect of broadband into these two components.

gains from broadband internet. First we decompose the welfare gains into nominal wage and price index components. Second, we decompose those gains into import, export and residual channels.[37] Results from this decompostion are reported in Table 4.5.

Table 4.5: Welfare changes for different values of θ and σ

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Parameters		**Endogenous wage**						**Exogenous wage**	
		log difference in			% contributions			log difference in	
θ	σ	nominal wage	price index	welfare	exports	imports	residual	price index	import chan.
2.70	3.50	2.43	-0.58	3.01	14.25	37.87	47.88	-1.85	-0.75
2.70	2.20	1.82	-0.69	2.51	22.34	28.51	49.16	-1.85	-0.57
2.70	5.00	2.93	-0.48	3.41	10.12	43.05	46.83	-1.85	-0.84
1.79	2.79	2.96	-1.10	4.06	12.05	22.93	65.02	-2.79	-0.68
4.54	3.85	1.69	-0.24	1.93	20.96	63.53	15.50	-1.10	-0.78

NOTES: : This table presents the log-change $\times 100$ in nominal wage (column 3), consumer price (column 4) and overall welfare (column 5) induced by broadband expansion to all cities simultaneously. Column (6) and (7) display the contribution in percents of broadband through exports and through imports. The last column (8) presents the contribution of residual factors (through gains in TFP). Each row corresponds to a different combination of values for the key parameters θ and σ. Our baseline in row (1) correspond to θ estimate from Fajgelbaum et al. (2019). The σ estimate comes from Berthou and Fontagné (2016). Columns (9) and (10) recall results from the analysis ignoring wage adjustment as presented in Section 7.1. Note that the import-channel presented in Column (10) depends indirectly in σ through the calibration of γ (see Table 4.7 for more details on the calibration).

Table 4.5 provides results of our welfare computations for five different combinations of values assigned to θ and σ, based on estimates from the literature. Columns (3) to (5) report changes in nominal wage, price index and overall welfare predicted by our model under those different parameterizations. National expansion of broadband internet results in a consumer surplus gain between 2 and 4%, with the change in nominal wage contributing the lion's share of the overall impact.

The first row constitutes our benchmark case. Our preferred estimate of σ is the one using firm-level intensive margin trade elasticity by Berthou and Fontagné (2016), while θ is taken from Fajgelbaum et al. (2019). Welfare gains are slightly above 3% in that case. Varying the value of σ in the second and third rows shows that welfare gains are positively associated with the demand elasticity. The fourth row takes Antras et al. (2017) as a base for θ and σ estimates. This increases substantially welfare gains, which is mostly due to a lower value of θ. Comparing the fifth row with the first one, we confirm that welfare gains fall with θ (which measures how homogeneous firm-city productivity is in the model).

It is also possible to compute which share of the welfare gains are directly caused by increased imports and exports (columns 6 to 8). We see that exports contribute only modestly to the overall increase in welfare, while the import channel and the residual gains (coming through TFP increases) are more comparable in magnitude. Columns (9) and (10) provide results for

[37] The formula for this decomposition can be found in equation (4.49).

the version of the model ignoring wage adjustment as presented in Section 7.1. A first remark from the variation over rows of the exogenous wage case is that the change in price index due to broadband is unaffected by the value of σ, (which is consistent with the left-hand-side of equation 4.8).[38] A second remark can be made from a comparison of welfare effects with and without allowing for broadband to have an effect on exports and nominal wages. When the model includes endogenous wages and exports, welfare gains from broadband increase substantially. This is true across all values of parameters proposed in Table 4.5: the increase in nominal wages largely dominates the lesser gains on the price index brought by the expansion of broadband internet.

8. Conclusion

We find broadband expansion to have substantially changed the importing patterns of French small and medium-sized firms. The affected firms increased their overall imports by around a quarter. This was driven by a rise in both the number of source countries and the count of products imported. The rise in imported value exceeded the growth of turnover, resulting in a large increase in import intensity (the imports over sales ratio goes up by about 15%).

The first implication of our findings relates to the debate regarding trade vs technology as separate explanations for labor market outcomes. For instance, Autor et al. (2015) argue that trade shocks had more impact than technological shocks on the decline of US manufacturing employment over the 2000's. Our work suggests that it is hard to empirically disentangle the effects of trade from those of technology: a specific instance of technological change, namely broadband internet, has contributed to the trade shocks that this literature typically takes as given. This interaction is quantitatively large, since we estimate that the rise in import penetration in France over the 1997-2007 period would have been 16% smaller without broadband internet.

A second implication is that the positive effect of broadband internet on productivity (Akerman et al., 2015; Ciapanna and Colonna, 2019) and welfare could be partly driven by the rise in the value and the variety of imported inputs. Under our preferred calibration, our event study estimates suggest that broadband internet resulted in a reduction of 1.7% in unit costs. The part of those gains that is attributable to enhanced access to foreign inputs (the import channel) is close to 25% of the total.

Trade and technology affect consumer welfare not only through prices but also through

[38] The fact that the import channel *does* depends on σ comes from the fact that in the model $1-\gamma = \frac{\text{expenditure on intermediates}}{\text{sales}} \times \frac{\sigma-1}{\sigma}$.

nominal earnings (Borusyak and Jaravel, 2018). It is plausible, that broadband internet induced labor market effects that were differentiated depending on whether firms engaged in international sourcing or not (Koren et al., 2019). Estimating the earning impact of fast internet by skill level and extending the theoretical model to account for richer interactions between purchase of foreign inputs and the skill-content of labor demand appears like an interesting avenue for future research.

References

Aghion, P., Bergeaud, A., Lequien, M., and Melitz, M. J. (2018). The impact of exports on innovation: Theory and evidence. Technical report, National Bureau of Economic Research.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, 130(4):1781–1824.

Akerman, A., Leuven, E., and Mogstad, M. (2018). Information frictions, internet and the relationship between distance and trade. *mimeo*, page 48.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Amiti, M. and Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review*, 97(5):1611–1638.

Amiti, M., Redding, S. J., and Weinstein, D. (2019). The impact of the 2018 trade war on us prices and welfare. Technical report, National Bureau of Economic Research.

Andrews, D., Nicoletti, G., and Timiliotis, C. (2018). Going digital: What determines technology diffusion among firms? Technical report.

Antras, P., Fort, T. C., and Tintelnot, F. (2017). The margins of global sourcing: Theory and evidence from us firms. *American Economic Review*, 107(9):2514–64.

Arcep (2002). L'accés haut débit via l'adsl : historique des décisions de l'art et de l'arcep. Technical report.

Arcep (2016). Etude sur les équipements et usages des pme et eti. Technical report.

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1):94–130.

Autor, D., Dorn, D., Hanson, G. H., Pisano, G., and Shu, P. (2016a). Foreign competition and domestic innovation: Evidence from us patents. Technical report, National Bureau of Economic Research.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics*, 21(1):1–42.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, 103(6):2121–68.

Autor, D. H., Dorn, D., and Hanson, G. H. (2015). Untangling trade and technology: Evidence from local labour markets. *The Economic Journal*, 125(584):621–646.

Autor, D. H., Dorn, D., and Hanson, G. H. (2016b). The china shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics*, 8:205–240.

Badré, D. (2007). Tarification de l'usage d'internet : Question écrite n 10755. *Senat, 11eme législature.*

Baldwin, R. (2016). *The great convergence*. Harvard University Press.

Baldwin, R. E. (2009). *The great trade collapse: Causes, consequences and prospects*. Cepr.

Barbero, J. and Rodriguez-Crespo, E. (2018). The effect of broadband on eu trade: a regional spatial approach. *The World Economy*.

Bas, M. and Strauss-Kahn, V. (2014). Does importing more inputs raise exports? firm-level evidence from france. *Review of World Economics*, 150(2):241–275.

Berlingieri, G., Breinlich, H., and Dhingra, S. (2018). The impact of trade agreements on consumer welfare—evidence from the eu common external trade policy. *Journal of the European Economic Association*, 16(6):1881–1928.

Berthou, A. and Fontagné, L. (2016). Variable trade costs, composition effects and the intensive margin of trade. *The World Economy*, 39(1):54–71.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275.

Bertschek, I., Wolfgang, B., Kai, H., Benedikt, K., and Niebel, T. (2015). The economic impacts of broadband internet: A survey. *Review of Network Economics*, 14(4):201–227.

Blaum, J., Lelarge, C., and Peters, M. (2018). The gains from input trade with heterogeneous importers. *American Economic Journal: Macroeconomics*, 10(4):77–127.

Bloom, N., Draca, M., and Van Reenen, J. (2016). Trade induced technical change? the impact of chinese imports on innovation, IT and productivity. *The Review of Economic Studies*, 83(1):87–117.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs with an application to the estimation of the marginal propensity to consume. *mimeo*, page 33.

Borusyak, K. and Jaravel, X. (2018). The distributional effects of trade: Theory and evidence from the united states. *Available at SSRN 3269579*.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, 82(1):1–44.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, 1(1):451–488.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American economic review*, 99(4):1145–77.

Ciapanna, E. and Colonna, F. (2019). Is your broadband really broad? internet speed, labour demand and productivity outcomes: Evidence from italian firms. *mimeo*.

Clarke, G. R. and Wallsten, S. J. (2006). Has the internet increased trade? developed and developing country evidence. *Economic Inquiry*, 44(3):465–484.

Cohen, D., Garibaldi, P., and Scarpetta, S. (2004). *The ICT revolution: Productivity differences and the digital divide*. Oxford University Press.

Dauth, W., Findeisen, S., and Suedekum, J. (2014). The rise of the east and the far east: German labor markets and trade integration. *Journal of the European Economic Association*, 12(6):1643–1675.

DeStefano, T., Kneller, R., and Timmis, J. (2018). Broadband infrastructure, ict use and firm performance: Evidence for uk firms. *Journal of Economic Behavior & Organization*, 155:110–139.

Eaton, J. and Kortum, S. (2001). Trade in capital goods. *European Economic Review*, 45(7):1195–1235.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: Evidence from french firms. *Econometrica*, 79(5):1453–1498.

Fajgelbaum, P. D., Morales, E., Suárez Serrato, J. C., and Zidar, O. (2019). State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376.

Feenstra, R. C. and Weinstein, D. E. (2017). Globalization, markups, and us welfare. *Journal of Political Economy*, 125(4):1040–1074.

Fort, T. C. (2017). Technology and production fragmentation: Domestic versus foreign sourcing. *The Review of Economic Studies*, 84(2):650–687.

Fort, T. C., Pierce, J. R., and Schott, P. K. (2018). New perspectives on the decline of us manufacturing employment. *Journal of Economic Perspectives*, 32(2):47–72.

Gopinath, G. and Neiman, B. (2014). Trade adjustment and productivity in large crises. *American Economic Review*, 104(3):793–831.

Gross, T., Notowidigdo, M. J., and Wang, J. (2018). The marginal propensity to consume over the business cycle.

Halpern, L., Koren, M., and Szeidl, A. (2015). Imported inputs and productivity. *American Economic Review*, 105(12):3660–3703.

Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica: Journal of the Econometric Society*, pages 541–561.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, volume 4, pages 131–195. Elsevier.

Hjort, J. and Poulsen, J. (2018). The arrival of fast internet and employment in africa. *American Economic Review*.

Houngbonon, G. V. and Liang, J. (2017). Broadband internet and income inequality.

Houngbonon, G. V. and Liang, J. (2018). The impact of broadband internet on employment in france. *Available at SSRN 3112182*.

Juhász, R. and Steinwender, C. (2018). Spinning the web: The impact of ict on trade in intermediates and technology diffusion. Technical report, National Bureau of Economic Research.

Kneller, R. and Timmis, J. (2016). Ict and exporting: The effects of broadband on the extensive margin of business service exports. *Review of International Economics*, 24(4):757–796.

Koren, M., Csillag, M., and Köllo, J. (2019). Machines and machinists: Importing skill-biased technology. *mimeo*.

Kovak, B. K. (2013). Regional effects of trade reform: What is the correct measure of liberalization? *American Economic Review*, 103(5):1960–76.

Le Gall, F. (2007). Adsl : France télécom a équipé tous les nra. Technical report.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. (2014). Market size, competition, and the product mix of exporters. *American Economic Review*, 104(2):495–536.

Muendler, M.-A. (2017). Trade, technology, and prosperity: An account of evidence from a labor-market perspective. Technical report, WTO Staff Working Paper.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics*, 122(2):569–600.

OECD (2004). ICT, E-Business And SMEs. *Promoting Entrepreneurship and Innovative SMEs in a Global Economy: Towards a More Responsible and Inclusive Globalisation.*

Portugal-Perez, A. and Wilson, J. S. (2012). Export performance and trade facilitation reform: Hard and soft infrastructure. *World development*, 40(7):1295–1307.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35.

Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of international Economics*, 62(1):53–82.

Sandoz, C. (2017). Input prices, allocation of resources and tfp growth: Evidence from chinese imports in france. *Job market paper, PSE.*

Schmidheiny, K. and Siegloch, S. (2019). On event study designs and distributed-lag models: Equivalence, generalization and practical implications. *CEPR Discussion Paper No. DP13477.*

Sénat (2002). Le bilan de la loi n 96-659 de réglementation des télécommunications. *Rapport d'information du Sénat*.

Shu, P. and Steinwender, C. (2019). The impact of trade liberalization on firm productivity and innovation. *Innovation Policy and the Economy*, 19(1):39–68.

Steinwender, C. (2018). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–96.

Suárez Serrato, J. C. and Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, 106(9):2582–2624.

Telecom, F. (2003). Conférence de presse du 10 juin 2003 "internet haut débit pour tous : France télécom s'engage". Technical report.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from india. *American Economic Journal: Applied Economics*, 2(4):1–41.

Topalova, P. and Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of india. *Review of economics and statistics*, 93(3):995–1009.

Tregouet, R. (2001). Question écrite *n*. 30844 (sénat, 01/02/2001). *Senat*.

A. Tables and figures

A.1. Baseline tables

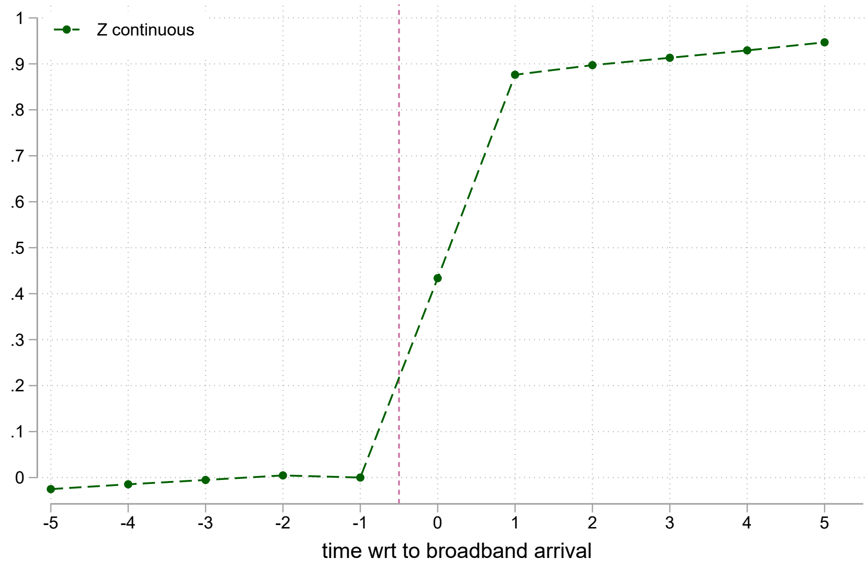
Table 4.6: Specification checks for main specification: ln(value of imports)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\{d = -5\}$	-0.010	-0.011	-0.015	-0.015	-0.012		
	(0.014)	(0.014)	(0.015)	(0.015)	(0.017)		
$\mathbb{1}\{d = -4\}$	-0.011	-0.012	-0.018	-0.018	-0.020		
	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)		
$\mathbb{1}\{d = -3\}$	-0.010	-0.018	-0.025*	-0.025	-0.026*		
	(0.013)	(0.014)	(0.015)	(0.015)	(0.016)		
$\mathbb{1}\{d = -2\}$	-0.011	-0.015	-0.020	-0.020	-0.017		
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)		
$\mathbb{1}\{d = 0\}$	0.003	0.008	0.016	0.016	0.023*	0.017	0.027
	(0.012)	(0.013)	(0.013)	(0.013)	(0.014)	(0.016)	(0.016)
$\mathbb{1}\{d = +1\}$	0.030*	0.041**	0.060***	0.059***	0.064***	0.050**	0.069***
	(0.018)	(0.019)	(0.019)	(0.019)	(0.020)	(0.022)	(0.022)
$\mathbb{1}\{d = +2\}$	0.076***	0.094***	0.125***	0.123***	0.137***	0.102***	0.133***
	(0.024)	(0.025)	(0.028)	(0.029)	(0.030)	(0.028)	(0.030)
$\mathbb{1}\{d = +3\}$	0.085***	0.115***	0.163***	0.161***	0.167***	0.123***	0.170***
	(0.031)	(0.034)	(0.039)	(0.040)	(0.041)	(0.037)	(0.041)
$\mathbb{1}\{d = +4\}$	0.098**	0.144***	0.212***	0.208***	0.211***	0.152***	0.218***
	(0.038)	(0.039)	(0.048)	(0.050)	(0.049)	(0.043)	(0.050)
$\mathbb{1}\{d = +5\}$	0.102**	0.156***	0.249***	0.243***	0.242***	0.164***	0.255***
	(0.049)	(0.050)	(0.061)	(0.065)	(0.062)	(0.055)	(0.064)
R2	0.84	0.85	0.85	0.85	0.85	0.85	0.85
Year FE	Yes	No	No	No	No	No	No
Year x Dep. FE	No	Yes	Yes	Yes	No	Yes	Yes
Year x ZE FE	No	No	No	No	Yes	No	No
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Density controls	No	No	Yes	Yes	Yes	No	Yes
control_f Fiscal + Sector + Educ. controls	No	No	No	Yes	No	No	No
Spec.	D	D	D	D	D	SD	SD
N	96947	96947	96916	96916	96905	96947	96916

NOTES : This table presents event-study estimates based on Equations (4.2) (columns 1 to 5) and (4.3) (columns 6 and 7). Robust standard errors clustered at the département level are presented in brackets. * p < 0.1, ** p < 0.05, *** p < 0.01. Estimates are plotted in Figure 4.3, see the figure notes for more details on the controls.

A.2. Pseudo first-stage

Fig. 4.15. The evolution of continuous measure of broadband coverage ($\widetilde{Z}_{it}$) around the (discrete) year of the largest increase in $\widetilde{Z}_{it}$ (t_{0i}).



NOTES: This figure plots estimates from regression of $\widetilde{Z}_{it}$ on set of time to broadband expansion dummies with year fixed effect. The time at -1 is normalized to 0. 95% confidence intervals are displayed (but graphically invisible due to their small size). The sample include all cities with a positive trade flow (import).

B. Conceptual framework for quantification: full derivations

In this section, we set up a simple partial equilibrium model of firm-level imports that allows to quantify the magnitude of the import channel in terms of consumer welfare. We have empirically estimated the effect of BI on imports and sales. We show below that, in a model featuring monopolistic competition and a CES demand system (MC-CES henceforth), we can recover the total effect of broadband on firm-level productivity (defined as the inverse of its unit-cost) from estimates of its impact on firm-level sales.

However, because broadband internet affects firm-performance through channels other than trade, it cannot be considered a valid instrument for imports when explaining unit costs. This is akin to a failure of the exclusion restriction leaving us with an under-identification problem: even if we know the total effect of BI on imports and productivity, we cannot isolate the contribution of broadband-induced imports to the overall effect of broadband internet on firm-level unit cost. To solve this problem, we add further structure on the production process and import environment, building on Blaum et al. (2018). In this simple theoretical framework, we show that a firm's observed domestic share of inputs (i.e. 1 minus its import share) encapsulates all of the effects of importing activities on its unit cost. Our main result is an expression for the overall effect of broadband internet on the price index and the contribution of the import channel as a function of *(i)* firm-level effects of broadband internet on sales and the domestic share of inputs and *(ii)* structural parameters pertaining to preferences and technology. We use our reduced-form estimates and calibrate the structural parameters based on existing work and/or aggregate data in order to quantify the effect of broadband internet on consumer welfare and the contribution of the import-channel.

We model France as a small open economy composed of N cities indexed by i. We assume that trade costs within France across cities are negligible. In our baseline model, a fixed mass of firms, normalized to be of measure one, pick cities in order to maximize profits. A firm j's productivity depends on local productive amenities i and an idiosyncratic productivity draw. The labor market is competitive. We assume that local prices and consumption amenities do not vary across areas. There is a mass of workers of measure $\mathbf{L}$ who are perfectly mobile so that wages are equalized across cities.

This section is structured as follows. In subsection B.1, we present the demand side of the economy. In subsection B.2, we present the firm's location choice and its implications for the distribution of economic activity across cities. In subsection B.3, we detail the production

function and import environment facing firms. We derive a sufficiency result in subsection B.4, which allows us to infer unit cost reductions due to trade from change in the input share of foreign goods. In subsection B.5, we show how we can express changes in consumer prices due to the import-channel as a function of our event-study estimates and present a first set of welfare implications of broadband on consumer welfare. As will be made clear below, we derive the implications of broadband for consumer welfare through the price index, (i) assuming away the impact of broadband on exports and (ii) considering that the price of the domestic intermediate goods is fixed and exogenous. We relax each of these assumptions in the last sections. We revisit the impact of broadband on exports and its implications on consumer welfare through the nominal wage channel in subsection B.6.

B.1. Demand and market structure

The representative consumer (we do restrict attention to a single country of consumption, France in this case) has a classical Constant Elasticity of Substitution (CES, denoted σ) utility function over the set Ω of available varieties of a final good:

$$U = \left(\int_{j \in \Omega} c_j^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}},$$

where c_i is the quantity consumed of variety i.

Those varieties are produced domestically under monopolistic competition. The CES demand function aggregated at the level of the nation yields sales x_i for a variety (also a firm) j:

$$x_j \equiv c_j p_j = p_j^{1-\sigma} E P^{\sigma-1}, \tag{4.9}$$

where E and P respectively represent total expenditure and the ideal CES price index in this economy and p_j represents the price of variety j. We consider a case where each firm will produce only a single variety which implies that Ω is of measure 1.

B.2. Firm location choice and city-level sales

Combined with monopolistic competition and constant marginal costs, the CES demand system implies a fixed markup over cost $\left(\frac{\sigma}{\sigma-1}\right)$, and therefore the values of sales writes as:

$$x_{ij} = u_{ij}^{1-\sigma} \times S \tag{4.10}$$

where u_{ij} refers to the unit cost of firm / variety j when choosing city i and $S \equiv EP^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}$ is a market wide demand-shifter. Given our assumption that trade costs within France are null this aggregate demand term is not i specific.

We assume that the productivity of firm j in city i is equal to a city-level component divided by an idiosyncratic productivity that is specific to the pair of the firm j when locating in city i. Formally we have:

$$u_{ij} = u_i / b_{ij}, \tag{4.11}$$

where the city-wide unit cost common to all firms locating in i is denoted u_i and b_{ij} refers to the idiosyncratic productivity draw. We assume b_{ij} follows a Fréchet Type 2 distribution with shape parameter θ (Fajgelbaum et al., 2019). Given constant return to scale and the MC-CES combination, profits are proportional to sales. Denoting π_{ij} the profit realized by firm j in city i, we can write:

$$\pi_{ij}^V = \underbrace{\frac{S}{\sigma} u_i^{1-\sigma}}_{\equiv v_i} \times b_{ij}^{\sigma-1} = v_i b_{ij}^{\sigma-1}. \tag{4.12}$$

Given our definition of v_i as the non-stochastic component of profit in city i, and the distributional assumption regarding b_{ij}, we can write the probability that a firm j choose city i as:

$$\mathbb{P}_i \equiv \text{Prob}(i = \text{argmax}_{i'}\ \pi_{i'j}^V) = \left(\frac{v_i}{v}\right)^{\frac{\theta}{\sigma-1}},$$

where $v \equiv \left(\sum_i v_i^{\frac{\theta}{\sigma-1}}\right)^{\frac{\sigma-1}{\theta}}$. We can also express this probability as a function of the common component of the unit cost in i:

$$\mathbb{P}_i = \frac{u_i^{-\theta}}{\sum_i u_i^{-\theta}}. \tag{4.13}$$

The aggregate sales by firms located in city i depend on the share of firms locating in i times

the average sale per firm conditional on choosing city i:

$$\begin{aligned}
x_i &= \mathbb{P}_i \times u_i^{1-\sigma} \times S \times \mathbb{E}(b_{ij}^{\sigma-1} | i = \text{argmax}_{i'}\ \pi^V_{i'j}) \\
&= \mathbb{P}_i^{\frac{\theta+1-\sigma}{\theta}} u_i^{1-\sigma} \times S \times \Gamma\left(\frac{\theta-\sigma+1}{\theta}\right) \\
&= u_i^{-\theta} \times \underbrace{\left(\sum_i u_i^{-\theta}\right)^{-\frac{\theta+1-\sigma}{\theta}}}_{\equiv u} \times S \times \Gamma\left(\frac{\theta-\sigma+1}{\theta}\right) \\
&= u_i^{-\theta} \times u \cdot S \cdot \Gamma\left(\frac{\theta-\sigma+1}{\theta}\right), \qquad (4.14)
\end{aligned}$$

where we go from line 1 to 2 by using a result contained in the classical work of Hanemann (1984) on discrete/continuous models,[39] and then from line 2 to line 3 by using (4.13).

Impact of broadband internet on sales and unit cost. Imposing this simple structure, we aim to infer the effect of broadband internet on the unit cost for a given value of θ. Our estimate of the effect of BI on total *domestic* sales d years after broadband internet expansion is denoted as $\widehat{\beta}_d^{\text{dsales}}$. We will use our estimate of the long-run impact (after 5 years), β_5^{dsales}. Using the sales equation (4.14), and assuming that the elasticity of substitution of the final good is left unchanged by the arrival of the internet, we can write:

$$\ln\left(\frac{x_i(1)}{x_i(0)}\right) = -\theta \ln\left(\frac{u_i(1)}{u_i(0)}\right) + \ln\left(\frac{S(1)}{S(0)}\frac{u(1)}{u(0)}\right) = \beta_5^{\text{dsales}}, \qquad (4.15)$$

where $X_i(1)$ and $X_i(0)$ are the values taken by X_i when $\text{BI}_{c(i)} = 1$ and $\text{BI}_i = 0$ respectively. The small size of individual cities i implies that each of them should have a negligible impact on the overall price index P as well as on aggregate spending E, implying $\frac{S(1)}{S(0)} \approx 1$ and $\frac{u(1)}{u(0)} \approx 1$. As a consequence, equation (4.15) implies that the effect of access to BI on domestic sales can be attributed to a fall in costs in the city where firm i is located and produces. Let us denote $\psi_i \equiv u_i(1)/u_i(0)$. Assuming, consistently with our empirical analysis, that all cities experience the same treatment effect, i.e. $\psi_i = \psi, \forall i \in \Omega$, we obtain

$$\ln\left(\frac{u_i(1)}{u_i(0)}\right) = \ln\psi = -\beta_5^{\text{dsales}}/\theta. \qquad (4.16)$$

[39] Hanemann (1984) equation (3.15) can be used to establish that $\mathbb{E}(b_{ij}^{\sigma-1} | i = \text{argmax}_{i'}\ \pi^V_{i'j}) = \mathbb{P}_i^{\frac{1-\sigma}{\theta}} \times \Gamma\left(\frac{\theta-\sigma+1}{\theta}\right)$.

Combining our estimate of β_5^{dsales} with a value for θ (that one can obtain from the literature's estimates), equation (4.16) provides the total effect of BI on the reduction of unit costs.

Price index. We now turn to consumer surplus. In this single sector model, the price index writes as:

$$P = \left(\int_{i\in\Omega} p_i^{1-\sigma} di\right)^{\frac{1}{1-\sigma}} = \frac{\sigma}{\sigma-1}\left(\int_{i\in\Omega} u_i^{1-\sigma} di\right)^{\frac{1}{1-\sigma}}. \tag{4.17}$$

We denote $P(\mathbf{0})$ the price index when no firm/cities has access to broadband internet, i.e. a fraction 0 of them, and symmetrically $P(\mathbf{1})$ the price index when all firms/cities have access to broadband internet. We can thus write the change in the price index as:

$$\ln\left(\frac{P(\mathbf{1})}{P(\mathbf{0})}\right) = \frac{1}{1-\sigma}\ln\left(\int_{i\in\Omega} \psi^{1-\sigma} \times \left[\frac{u_i(0)^{1-\sigma}}{\int_{i\in\Omega} u_i(0)^{1-\sigma} di}\right] di\right) = \ln\psi. \tag{4.18}$$

Note that the $\frac{u_i(0)^{1-\sigma}}{\int_{i\in\Omega} u_i(0)^{1-\sigma} di}$ term is equal to the market share of firm i (in the absence of BI), which would weight the treatment in case of heterogeneity (firm-level or spatial ψ_i). Because of the constant treatment effect assumption, we obtain an estimate of the change in the price index directly from our firm-level estimate of the impact of BI on sales:

$$\ln\left(\frac{P(\mathbf{1})}{P(\mathbf{0})}\right) = \ln\psi = -\beta_5^{\mathbf{dsales}}/\theta. \tag{4.19}$$

B.3. Production function

We now specify the production function and the import environment, building on Blaum et al. (2018)'s setup. We assume that the technology we describe below is common to all firms located in city i. Accordingly we will refer to city and firm interchangeably in this section, although one must keep in mind that, within each city i, firms are heterogeneous in terms of their idiosyncratic draw b_{ij}. Each firm operates a Cobb-Douglas production function combining labor and intermediate goods, with the exponent on labor denoted as γ:

$$Y_i = A_i L_i^{\gamma} X_i^{1-\gamma} \tag{4.20}$$

where Y_i refers to total physical production, A_i is the total factor productivity of the firm, L_i is the effective quantity of labor, and X_i the effective quantity of intermediate inputs. The parameter γ determine the labor share in total costs.

Intermediates can be either domestic or foreign and they are aggregated through a CES

function with elasticity of substitution higher than unity:

$$X_i = \left(\alpha_D X_{i,D}^{\frac{\varepsilon-1}{\varepsilon}} + \alpha_F X_{i,F}^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{4.21}$$

where $X_{i,D}$ is the quantity of domestic inputs while $X_{i,F}$ refers to the quantity of imported inputs. The parameter $\varepsilon > 1$ determines the elasticity of substitution between foreign and domestic inputs. The terms α_D and α_F parameterize the relative quality of domestic versus foreign inputs. Finally, the foreign input $X_{i,F}$ itself aggregates imported varieties through a constant return to scale production function $G_i()$ which can be city-specific:

$$X_{i,F} = G_i(\{q_o X_{i,o}\}_{o\in\Omega_i}). \tag{4.22}$$

Foreign inputs $X_{i,o}$ are indexed according to their country of origin o. The term q_o parameterizes the relative efficiency of inputs from country o. Ω_i refers to the set of countries from which the firm sources its inputs (its *sourcing strategy*) and is potentially an endogenous variable (as in Antras et al., 2017, for instance).

B.4. Sufficiency result

In that setting, conditioning on the set of sourcing countries Ω_i, the unit-cost is given by:

$$u_i \equiv \left\{\min_{L,X} Lw + X_D p_{X_D} + \sum_{c\in\Omega_i} X_{F,c} p_{F,c}, \text{ s.t. } Y_i \geq 1\right\}$$

where $p_{F,o}$ is the price of input varieties imported from o. We can define the (ideal) import price index as the ratio of the spending on imports over the effective quantity of imports as defined in (4.22):

$$\mathbf{p}_i(\Omega_i) \equiv E_{i,F}/X_{i,F}$$

where $E_{i,F}$ denotes expenditure on foreign goods by firm i. Note that we leave, for now, the function $G_i()$ unspecified. However for most plausible aggregators, $\mathbf{p}_i(\Omega_i)$ will be declining in the cardinality of Ω_i ($|\Omega_i|$), meaning that the effective price index of imports will decrease in the number of sourcing countries. In the presence of fixed costs of importing that are paid per destination, there is a trade-off between higher payment towards origin-specific fixed costs of importing and the resulting reduction in the price index of imported goods which ultimately translates into lower overall unit cost. This can occur in Armington-type models where inputs are differentiated by country of origin, thus creating a love-for-variety effect (Halpern et al., 2015; Sandoz, 2017) or in models where expanding the set of sourcing

countries allows accessing cheaper quality-adjusted inputs (Antras et al., 2017).

Given the price index of imports $\mathbf{p}_i(\Omega_i)$, and the CES aggregation of foreign and domestic inputs, we can express the price index of all intermediates as:

$$Q_i(\Omega_i) = \left(\alpha_D^{\varepsilon} p_{X_D}^{1-\varepsilon} + \alpha_F^{\varepsilon} \mathbf{p}_i(\Omega_i)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}. \tag{4.23}$$

Note that we assume that trade cost within France are null. This implies that p_{X_D}, the price of the *domestic* intermediate goods is constant across locations and does not depend on whether a city has broadband or not. Given the Cobb-Douglas assumption at the upper level, it is straightforward to express unit cost as:

$$u_i = (1-\gamma)^{-(1-\gamma)} \gamma^{-\gamma} \times A_i^{-1} Q_i(\Omega_i)^{1-\gamma} w^{\gamma}. \tag{4.24}$$

Given the CES aggregation between domestic and foreign inputs, we can express the equilibrium share of domestic inputs as:

$$sd_i = \alpha_D^{\varepsilon} p_{X_D}^{1-\varepsilon} Q_i(\Omega_i)^{\varepsilon-1}, \tag{4.25}$$

where $sd_i \equiv p_{X_D} X_{D,i} / (X_{F,i}\mathbf{p}(\Omega_i) + p_{X_D} X_{D,i})$ is the fraction of firm spending on intermediates that goes to domestic inputs. We can relate unit costs and domestic share of intermediate inputs by isolating $Q_i(\Omega_i)$ in (4.25), then substituted into (4.24):

$$u_i = \tilde{A}_i^{-1} sd_i^{\frac{1-\gamma}{\varepsilon-1}} p_{X_D}^{1-\gamma} w^{\gamma} \tag{4.26}$$

where $\tilde{A}_i = A_i(\alpha_D)^{\frac{\varepsilon(1-\gamma)}{\varepsilon-1}}(1-\gamma)^{1-\gamma}\gamma^{\gamma}$. Equation (4.26) is essentially a sufficiency result, as first demonstrated by Blaum et al. (2018) (see their equation 6), stating that conditional on its domestic expenditure share (i.e. 1 minus the importing share), the rest of the import environment (including the global sourcing strategy Ω_i, prices p_{oi} or the technology through which foreign goods are agregated $G_i()$) does not affect the firm's unit cost. The domestic expenditure share conveys all the information from the inputs importing activity that is relevant for the unit cost.[40] Note that this sufficient statistic result does *not* depend on a specific market structure assumption regarding the final good. Opting for CES preferences with monopolistic competition as we do here is useful to obtain the one-to-one mapping from changes in unit cost to changes in consumer prices shown in section B.1.

[40]We present the sufficiency result in a very simplified import environment (Armington model with a continuum of homogeneous countries with homogeneous variables and fixed costs for importing) that allows for a simple closed-form solution in the Online Appendix B.

B.5. *Broadband expansion and the sufficiency result*

On top of the increased imports and other trade-related impacts, the expansion of broadband internet is likely to affect unit costs u_i through additional channels. It could for instance affect A_i directly.

Using equation (4.26) and maintaining the notation (0) and (1) for the potential outcomes taken without and with BI respectively, we obtain:

$$\underbrace{\ln\left(\frac{u_i(1)}{u_i(0)}\right)}_{\text{total effect}} = \ln\psi = \underbrace{\frac{1-\gamma}{\varepsilon-1}\ln\left(\frac{sd_i(1)}{sd_i(0)}\right)}_{\text{import channel}} - \underbrace{\ln\left(\frac{\tilde{A}_i(1)}{\tilde{A}_i(0)}\right)}_{\text{residual effect}}. \tag{4.27}$$

The import-channel is fully summarized by the first underbraced term $\frac{1-\gamma}{\varepsilon-1}\ln\left(\frac{sd_i(1)}{sd_i(0)}\right)$. The first equality comes from the constant treatment effect assumption that we maintain throughout.[41]

Mapping reduced-form estimates of BI into welfare effects. The overall contribution of broadband internet to the decline in unit cost can be expressed as:

$$\ln\left(\frac{u_i(1)}{u_i(0)}\right) = \ln\psi = -\frac{1}{\theta}\ln\left(\frac{x_i(1)}{x_i(0)}\right) = -\frac{\beta_5^{\mathbf{dsales}}}{\theta}, \tag{4.28}$$

where $\beta_5^{\mathbf{dsales}}$ refers to the 5-year horizon effect of BI on the log of sales. Given our estimate $\widehat{\beta}_5^{\mathbf{dsales}}$ and assuming $\theta = 2.7$ (based on Fajgelbaum et al. (2019)), we have:[42]

$$\text{Total price channel} = \ln\left(\frac{u_i(1)}{u_i(0)}\right) = -\frac{\beta_5^{\mathbf{dsales}}}{\theta} = -\frac{0.05}{2.7} = -1.85\%.$$

To locate this estimate in the literature, we note that the figure of -1.85% is about half the productivity effect estimated by Akerman et al. (2015) in Norway. They found that a 10% increase in broadband availability is associated with a 0.4% increase in productivity

[41] Note that we take γ here to be a policy invariant parameter. This assumption appears at odds with the empirical findings of Fort (2017) who shows that ICT adoption in the US increases the probability of both domestic and foreign sourcing by firms. Her results would suggest a positive effect of broadband expansion on $1-\gamma$. However, in our application, we do not find any effect of broadband internet on the empirical counterpart of $1-\gamma$. The point estimate after 5 periods is -0.004 with 95% confidence interval: $(0.0036, -0.0116)$. Results are displayed Figure 4.22 of the online appendix. It seems therefore warranted to model γ as a fixed (i.e. treatment invariant) parameter.

[42] This value corresponds to the mean price elasticity of sales in gravity equations (estimated with country fixed-effects) reported in the meta-analysis by Head and Mayer (2014) based on 32 papers and 744 estimates.

(as measured by overall production holding inputs constant), suggesting a 4% increase when broadband availability goes from 0 to 100%.

The impact of BI access that goes through the *import channel* (via an increased access to foreign intermediates), can be expressed as:

$$\text{Import Price Channel} = \frac{1-\gamma}{\varepsilon - 1} \ln\left(\frac{sd_i(1)}{sd_i(0)}\right) = \frac{1-\gamma}{\varepsilon - 1} \times \beta_5^{\mathbf{sf}} \times m, \tag{4.29}$$

where $\beta_5^{\mathbf{sf}}$ refers to the 5-year horizon effect of BI on the log of the foreign share $sf \equiv 1 - sd$, and $m = sf/sd$.[43]

For the value of ε, we rely on Blaum et al. (2018)—whose estimation is also carried out on French data for a comparable period (2000-2006). For the overall economy we retain $\varepsilon = 2.3$. The model implies: $1 - \gamma = \frac{\text{expenditure on intermediates}}{\text{sales}} \times \frac{\sigma - 1}{\sigma}$. Using the sample average of the ratio of spending on intermediates over sales and $\sigma = 3.5$ yields $1 - \gamma = 0.47$. As a consequence, the import channel of the effect of BI writes as:

$$\begin{aligned} \text{Import Price Channel} &= \frac{1-\gamma}{\varepsilon - 1} \times (-m \times \beta_5^{\mathbf{sf}}) \\ &= -0.75\% \quad \text{including capital goods} \\ &= -0.58\% \quad \text{excluding capital goods.} \end{aligned}$$

Berlingieri et al. (2018) estimate in a recent paper that trade agreements implemented by the EU led to an overall decline of 0.24% in the consumer price index over a 20 year period (1993-2013). Therefore, while the import channel of broadband internet access might at first seem moderate, it is large compared to other policy experiments and accounts for about 33% of the consumer gains from broadband internet.

B.6. Extension: Exports and endogenous wage

Until this point, we have assumed that firms only sell on the domestic market. We have also assumed that broadband expansion, even when applied to the entire economy, is too small to have a direct impact on wages.

In this subsection, we relax those assumptions by introducing exports in our model. Our extension to allow for exports has no impact on the welfare effects that go through the price

[43] We use the approximation $\ln\left(\frac{sd_i(1)}{sd_i(0)}\right) \approx \frac{sd_i(1) - sd_i(0)}{sd_i(0)}$ and the definition of foreign share $sf = 1 - sd$ to write: $\ln\left(\frac{sd_i(1)}{sd_i(0)}\right) \approx -\frac{sf_i(1)}{sd_i(0)} \ln\left(\frac{sf_i(1)}{sf_i(0)}\right)$.

level of the import-channel that we derive above, but it introduces a new channel through which broadband impact consumer's welfare: higher nominal wages.

In each foreign market k, consumers are assumed to have the same elasticity of demand as the domestic consumers (σ). They are endowed with a specific income E_j and face a price index $P_k^{\sigma-1}$. Exporting from city i to country k entails an iceberg cost τ_{ik}. We do not model selection into destinations: there is no fixed cost of exporting and firms in all cities are supposed to export some positive amount to all k.

Overall exports by firm j in city i towards d writes as:

$$x_{ijk} = (p_{ij}\tau_{ik})^{1-\sigma} E_k P_k^{\sigma-1} = u_{ij}^{1-\sigma} \tau_{ik}^{1-\sigma} \times S_k, \tag{4.30}$$

in which the demand shifter S is now destination specific: $S_k \equiv \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} E_k P_k^{\sigma-1}$. Denoting the domestic market with D, we can write the sum of total exports from city i as:

$$x_{ij}^X = u_{ij}^{1-\sigma} \left(\sum_{k \neq D} \tau_{ik}^{1-\sigma} \times S_k \right). \tag{4.31}$$

Domestic sales (for which the trade cost is assumed to be nil) are given by:

$$x_{ij}^D = u_{ij}^{1-\sigma} \times S_D, \tag{4.32}$$

where the domestic demand shifter is $S_D = \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} E_D P_D^{\sigma-1}$.

Let us define $\Phi_i = \sum_{k \neq D} \tau_{ik}^{1-\sigma} \times S_k$ as the trade cost-weighted average of foreign demand, i.e. an export market potential (also called foreign market access or foreign market potential in the literature that followed Redding and Venables (2004)). Overall market potential also include domestic demand and is denoted with $\Psi_i = \Phi_i + S_D$.[44] Overall sales by firm j in city i, defined as $x_{ij} = x_{ij}^D + x_{ij}^X$,

$$x_{ij} = u_{ij} \times \Psi_i = u_i^{1-\sigma} b_{ij}^{\sigma-1} \times \Psi_i, \tag{4.33}$$

Variation in market potential across cities will affect entry at the city level. Profit by firm j in city i writes as:

$$\pi_{ij}^V = b_{ij}^{\sigma-1} \times \frac{1}{\sigma} \Psi_i u_i^{1-\sigma} = b_{ij}^{\sigma-1} v_i.$$

[44] Two points are worth noting: (1) the market potential related to foreign sales is city specific because we allow trade costs to be specific between each city i and destination k, (2) the domestic market potential is stable constant across cities. We make the assumption that the domestic market is fully integrated and trade costs across cities within France are null and therefore will not be affected by the introduction of broadband.

Recalling that the random part of productivity b_{ij} is distributed Fréchet, we obtain the location choice probability as:

$$\mathbb{P}_i \equiv \text{Prob}(i = \text{argmax}_{i'} \pi^V_{i'j}) = \frac{\Psi_i^{\frac{\theta}{\sigma-1}} u_i^{-\theta}}{\sum_i \Psi_i^{\frac{\theta}{\sigma-1}} u_i^{-\theta}}.$$

The value of total sales conditional on choosing city i is:

$$\begin{aligned} x^* &\equiv \mathbb{E}(x_{ij} | i = \text{argmax}_{i'} \pi^v_{i'j}) \\ &= u_i^{1-\sigma} \Psi_i \mathbb{E}(b_{ij}^{\sigma-1} | i = \text{argmax}_{i'} \pi^v_{i'j}) \\ &= u_i^{1-\sigma} \Psi_i \mathbb{P}_i^{-\frac{\sigma-1}{\theta}} \Gamma\left(\frac{\theta - \sigma + 1}{\theta}\right) \\ &= v \times \Gamma\left(\frac{\theta - \sigma + 1}{\theta}\right), \end{aligned}$$

where we define v as $v = \left(\sum_{i=1}^N \Psi_i^{\frac{\theta}{\sigma-1}} u_i^{-\theta}\right)^{\frac{\sigma-1}{\theta}}$. This allows to compute overall sales by firms in city i as:

$$\begin{aligned} x_i = \mathbb{P}_i x^* &= \mathbb{P}_i \times v \times \Gamma\left(\frac{\theta - \sigma + 1}{\theta}\right) \\ &= u_i^{-\theta} \Psi_i^{\frac{\theta}{\sigma-1}} v^{1-\frac{\theta}{\sigma-1}} \kappa, \end{aligned} \tag{4.34}$$

where $\kappa \equiv \Gamma\left(\frac{\theta-\sigma+1}{\theta}\right)$. The main difference with the city-level sales equation (4.14) is now that part of the demand is city-specific since Ψ_i varies across cities (through the export market potential Φ_i). Overall exports by firms locating in city i write as follows:

$$x_i^X = \mathbb{P}_i x_i^{X*} = u_i^{-\theta} \Psi_i^{\frac{\theta+1-\sigma}{\sigma-1}} \Phi_i v^{1-\frac{\theta}{\sigma-1}} \kappa.$$

From equation (4.34), we can see that exports as a share of sales is given by: $\theta^X = x_i^X / x_i = \Phi_i / \Psi_i$.

Equilibrium conditions on the labor market. Given the Cobb-Douglas production function, spending on labor is proportional to sales which are given in equation (4.34). Accordingly, local labor demand in city i writes as:

$$l_i = \gamma \frac{\sigma - 1}{\sigma} x_i / w = \gamma \frac{\sigma - 1}{\sigma} \kappa v^{1-\frac{\theta}{\sigma-1}} \times \frac{u_i^{-\theta} \Psi_i^{\frac{\theta}{\sigma-1}}}{w}. \tag{4.35}$$

We consider that there is a fixed endowment of labor of measure $\mathbf{L}$. The labor market clearing condition at the national level writes as:

$$\mathbf{L} = \sum_i l_i. \tag{4.36}$$

The impact of broadband Internet expansion on nominal wages. The definition of unit cost remain unchanged (see equations 4.24 and 4.26). For reasons of tractability, it is useful to express the change following the local introduction of broadband internet in terms of exact hat algebra where we will define the evolution of a variable x_i in city i as: $\widehat{x}_i = \frac{x_i(1)}{x_i(0)}$. When looking at aggregate variables, $\widehat{x} = \frac{x(\mathbf{1})}{x(\mathbf{0})}$ represents the change in outcome following the complete coverage of French cities with broadband. Consistent with our results which suggests fairly small level heterogeneity across firms, we assume constant treatment effects on productivity $\widehat{A}$, cost of intermediates $\widehat{Q}$ through trade, and export market potential $\widehat{\Phi}$.

A last assumption is that p_{X_D}, the price of the *domestic* intermediate goods is constant across locations and does not depend on whether a city has broadband or not. We further assume that this price is exogenous in general equilibrium as well.[45]

Using the definition of the unit cost of production in city i in equation (4.24), the change in the unit cost of production writes as:

$$\widehat{u}_i = \widehat{A}_i^{-1} \widehat{Q}_i^{1-\gamma} \widehat{w}_i^{\gamma} = \widehat{A}^{-1} \widehat{Q}^{1-\gamma} \widehat{w}^{\gamma} = \widehat{u}, \tag{4.37}$$

where the second equality stems from our assumption of constant treatment effect and the assumption that wages are determined by a national market clearing conditions—implying that $\widehat{w}_i = \widehat{w}$. Based on the definition of v, we obtain: $\widehat{v} = \widehat{u}^{1-\sigma} \widehat{\Psi}$.

Let us first consider changes in local labor demand l_i following broadband expansion, using (4.35), together with the definition of v and u:

$$\begin{aligned} \widehat{l}_i &= \widehat{u}_i^{-\theta} \widehat{\Psi}_i^{\frac{\theta}{\sigma-1}} \widehat{v}^{\frac{\sigma-1-\theta}{\sigma-1}} \widehat{w}^{-1} \\ &= \widehat{A}^{\sigma-1} \widehat{Q}^{(1-\sigma)(1-\gamma)} \widehat{\Psi} \widehat{w}^{\gamma(1-\sigma)-1} = \widehat{l}, \end{aligned} \tag{4.38}$$

[45] We relax this assumption in Online appendix section E.1 following a roundabout model of production (Caliendo and Parro, 2015) and show that the effects are scaled up by a factor $1/\gamma$ without affecting the relative contribution of the different channels.

Considering the national labor market clearing condition we have:

$$
\begin{aligned}
\widehat{\mathbf{L}} = 1 &= \sum_i s_i^L \widehat{l}_i, \quad \text{where } s_i^L = \frac{l_i(0)}{\sum_i l_i(0)} \\
&= \sum_i s_i^L \widehat{l} = \widehat{l}.
\end{aligned} \tag{4.39}
$$

Given our results on $\widehat{l}$ in Equation (4.38), we have:

$$
1 = \widehat{A}^{\sigma-1}\widehat{Q}^{(1-\sigma)(1-\gamma)}\widehat{\Psi}\widehat{w}^{\gamma(1-\sigma)-1}, \tag{4.40}
$$

Denoting $\lambda_D = 1 - \theta^X$, we can write $\widehat{\Psi} = \lambda_D 1 + (1-\lambda_D)\widehat{\Phi}$, and then express the change in wage as a function of the change in the market potential due to the impact of broadband on exports:

$$
\widehat{w} = \widehat{A}^{\frac{\sigma-1}{\gamma(\sigma-1)+1}}\widehat{Q}^{\frac{(1-\sigma)(1-\gamma)}{\gamma(\sigma-1)+1}} \times \left(\lambda_D 1 + (1-\lambda_D)\widehat{\Phi}\right)^{\frac{1}{\gamma(\sigma-1)+1}}. \tag{4.41}
$$

There are two parts in this expression for how wages are affected by BI. First, the direct effect on TFP (A) and costs of inputs (through imports) Q. The second channel is the increased market potential reflecting easier exports.[46] Using equation (4.37), we can write $\widehat{u}|_w \equiv \frac{u_i(1)}{u_i(0)}|_w = \widehat{A}^{-1}\widehat{Q}^{1-\gamma}$, and re-express (4.41) as

$$
\widehat{w} = (\widehat{u}|_w)^{\frac{1-\sigma}{\gamma(\sigma-1)+1}} \times \left(\lambda_D 1 + (1-\lambda_D)\widehat{\Phi}\right)^{\frac{1}{\gamma(\sigma-1)+1}}. \tag{4.42}
$$

We now have the needed elements to quantify the effect of broadband expansion on nominal wages. Taking logs of (4.42), we immediately obtain two separable effects:

1. A **Unit Cost Wage Channel**:

$$
\ln \widehat{w}^{\text{uc}} \equiv \frac{1-\sigma}{\gamma(\sigma-1)+1} \ln \widehat{u}|_w. \tag{4.43}
$$

2. An **Export Wage Channel**:

$$
\ln \widehat{w}^X \equiv \frac{1}{\gamma(\sigma-1)+1} \ln\left(\lambda_D + (1-\lambda_D)\widehat{\Phi}\right). \tag{4.44}
$$

The first comes from the productivity gains brought by broadband internet, while the second

[46] Note that in the derivation of this expression for the change in nominal wage, we neglect the impact of rising wages ($\widehat{w} > 1$) on the domestic demand shifter S_D.

is due to increased export opportunities.

Identification. The identification of the export wage channel in equation (4.44) requires information on $\widehat{\Phi}$, s the change in foreign market potential induced by broadband expansion. It can be retrieved from our event-study estimates. Noting that the export share of sales s^X is equal to $s_i^X = \Phi_i/\Psi_i$, the change in the export share writes as:

$$\widehat{s}^X = \widehat{\Phi}\widehat{\Psi}^{-1} = \frac{\widehat{\Phi}}{\lambda_D + (1-\lambda_D)\widehat{\Phi}} \Rightarrow \widehat{\Phi} = \frac{\lambda_D \widehat{s}^X}{1-\widehat{s}^X(1-\lambda_D)}.$$

Denoting $\beta_{s^X} = \ln(s^X(1)/s^X(0)) = \ln \widehat{s}^X$, we can express the export channel as:

$$\ln \widehat{w}^X = \frac{1}{\gamma(\sigma-1)+1} \ln\left(\lambda_D + (1-\lambda_D)\frac{\lambda_D \exp(\beta_{s^X})}{1-\exp(\beta_{s^X})(1-\lambda_D)}\right). \tag{4.45}$$

Based on derivations presented in section B.1—see equation (4.16)—we can express the average reduction in unit cost wage conditional on wage $\ln(\widehat{u}|_w)$, i.e. in partial equilibrium, as a function of the impact of broadband on local sales and the scale parameter of the Fréchet distribution. We can therefore identify the channel as:

$$\ln \widehat{w}^{\text{uc}} = \frac{1-\sigma}{\gamma(\sigma-1)+1} \times \left(-\frac{\beta_5^{\textbf{dsales}}}{\theta}\right). \tag{4.46}$$

The last needed elements for our quantification of $\ln \widehat{w}^X$ and $\ln \widehat{w}^{\text{uc}}$ are i) the average domestic share of sales λ_D, ii) the three structural parameters θ, γ and σ. The source and values of those parameters and all other needed elements for the quantification exercise are summarized in Table 4.7.

Welfare implications. We can finally turn to the total effect of the full coverage of broadband internet on workers' welfare.

$$\Delta \ln W = \Delta \ln w - \Delta \ln P.$$

Combining equation (4.18) with (4.37), we obtain an expression for the log change in the consumer price index allowing for wage changes as:

$$\Delta \ln P = \ln(\widehat{u}) = \ln\left(\widehat{A}^{-1}\widehat{Q}^{1-\gamma}\widehat{w}^{\gamma}\right) = \ln \widehat{u}|_w + \gamma \ln \widehat{w}. \tag{4.47}$$

Table 4.7: Parameters and identification

Panel **a**. Calibrated parameters			
Parameter	Explanation	Source	Values
σ	elasticity of subst. products	Berthou and Fontagné (2016)	3.5
ε	elasticity of subst. inputs	Blaum et al. (2018)	2.3
θ	scale param. Fréchet prod shocks	Fajgelbaum et al. (2019)	2.7

Panel **b**. Moments that can directly be estimated with sample statistics			
Moments	Explanation	Formula	Values
γ	labor share in CB function	$1-\gamma = \frac{\text{intermediates}}{\text{sales}} \times \frac{\sigma-1}{\sigma}$	0.47
λ_D	share of exports in total sales	NA	0.94
m	share of foreign over dom. inputs	NA	0.13

Panel **c**. Identification of partial effects from event-study estimates			
Partial effect	– as a function of estimable param.	Empirical implementation	Values
$\ln \widehat{Q}$	$-\frac{1}{1-\varepsilon}\ln(m \times \ln \widehat{\text{sf}})$	$-\frac{1}{1-\varepsilon}\ln(m \times \beta^{\mathbf{sf}})$	$\beta^{\mathbf{sf}} = 0.16$
$\ln \widehat{u}\|_w$	$-\frac{1}{\theta}\ln \widehat{x}$	$-\frac{1}{\theta}\beta^{\mathbf{dsales}}$	$\beta^{\mathbf{dsales}} = 0.05$
$\ln \widehat{\Phi}$	$\ln \frac{\lambda_D \widehat{s}^X}{1-\widehat{s}^X(1-\lambda_D)}$	$\ln \frac{\lambda_D \exp(\beta^{\mathbf{sx}})}{1-\exp(\beta^{\mathbf{sx}})(1-\lambda_D)}$	$\beta^{\mathbf{sx}} = 0.14$

NOTES: This table presents the parameters and reduced-form effects used to perform the welfare computation. The sample statistics used are computed for year 1999 (i.e. prior to any widespread broadband expansion).

Note that in equilibrium the impact of BI on the price index contains both the reduction of costs due to lower-cost inputs and a general equilibrium effect on wages with labor share γ.[47]

Table 4.8 provides results of our welfare computation for five different combinations of the two key elasticities obtained from the literature (θ and σ, reported in the first two columns). Columns (3) to (5) report changes in nominal wage, price index and overall welfare predicted by our model under those different parameterizations. National expansion of broadband internet results in a consumer surplus gain between 2 and 4.4%, with the change in nominal wage contributing the lion's share of the overall impact.

The first row constitutes our benchmark case. Parameter θ is obtained from the most comparable estimate of Fajgelbaum et al. (2019), which also model it as a firm-location productivity dispersion parameter. Our preferred estimate of σ is the one using firm-level intensive margin trade elasticity by Berthou and Fontagné (2016). Welfare gains are slightly

[47]Welfare can be restated as

$$\begin{aligned}\Delta \ln W &= (1-\gamma)(\ln \widehat{w}^X + \ln \widehat{w}^{\text{uc}}) - \left(-\frac{\beta_5^{\mathbf{dsales}}}{\theta}\right). \\ &= \ln \widehat{w}^X + \left(\frac{\sigma}{\gamma(\sigma-1)+1}\right)\left(\frac{\beta_5^{\mathbf{dsales}}}{\theta}\right),\end{aligned} \tag{4.48}$$

with $\sigma > 1$ ensuring that $\frac{\sigma}{\gamma(\sigma-1)+1} > 1$. This means that allowing for exports magnifies the partial equilibrium effect $\beta_5^{\mathbf{dsales}}/\theta$. It also adds a pure export effect on wages $\ln \widehat{w}^X$.

above 3% in that case. Varying the value of σ in the second and third rows shows that welfare gains are positively associated with the demand elasticity. The fourth row takes Antras et al. (2017) as a base for θ and σ estimates. This increases substantially welfare gains, which is mostly due to a lower value of θ. Comparing the fifth row with the first one, we confirm that welfare gains fall with θ (which measures how homogeneous firm-city productivity is in the model).

Table 4.8: Welfare changes for different values of θ and σ

Parameters		log difference in			% contribution of		
θ	σ	nominal wage	price index	welfare	exports	imports	residual
2.70	3.50	2.43	-0.58	3.01	14.25	37.87	47.88
2.70	2.20	1.82	-0.69	2.51	22.34	28.51	49.16
2.70	5.00	2.93	-0.48	3.41	10.12	43.05	46.83
1.79	2.79	2.96	-1.10	4.06	12.05	22.93	65.02
4.54	3.85	1.69	-0.24	1.93	20.96	63.53	15.50

NOTES: This table presents the log-change $\times 100$ in nominal wage (column 3), consumer price (column 4) and overall welfare (column 5) induced by broadband expansion to all cities simultaneously. Column (6) and (7) display the contribution in percents of broadband through exports and through imports. The last column (9) presents the contribution of residual factors (through gains in TFP). Each row corresponds to a different combination of values for the key parameters θ and σ. Our baseline in row (1) correspond to θ estimate from Fajgelbaum et al. (2019). The σ estimate comes from Berthou and Fontagné (2016). Note that γ varies with values of σ due to the calibration procedure (see Table 4.7 for details).

It is also possible to compute which share of the welfare gains are directly caused by increased imports and exports. Using (4.47) and the fact that $\ln \widehat{w} = \ln \widehat{w}^X + \ln \widehat{w}^{\text{uc}}$ in the welfare change equation, we obtain

$$\begin{aligned}\Delta \ln W &= \ln \widehat{w}^{\text{X}} - \frac{\sigma}{\gamma(\sigma-1)+1} \ln \widehat{u}|_w = \ln \widehat{w}^{\text{X}} - \frac{\sigma}{\gamma(\sigma-1)+1} \ln(\widehat{Q}^{1-\gamma}/\widehat{A}) \\ &= \ln \widehat{w}^{\text{X}} - \frac{\sigma(1-\gamma)}{[\gamma(\sigma-1)+1](\varepsilon-1)} \times m \ln \widehat{\text{sf}} + \frac{\sigma}{\gamma(\sigma-1)+1} \ln \widehat{A}, \end{aligned} \tag{4.49}$$

in which $m = \mathbf{sf}/\mathbf{sd}$ is the ratio of the shares of imported to domestic inputs, and $\ln \widehat{\text{sf}}$ is is the proportional change in the share of foreign inputs.

Coming back to results in Table 4.8, we see that exports contribute only modestly to the overall increase in welfare, while the import channel and the residual gains (coming for instance through TFP increases) are more comparable in magnitude.

ONLINE APPENDIX

C. Data appendix

C.1. Description of the datasets

Here we present in some more details the datatsets used in the analysis:

1. **DADS: employment at the establishment level**: The annual social data declaration (DADS) files contain information on each salaried contract for all firms of the competitive sector. Crucially, the place of work is documented at the city level. We use a version of this dataset that is aggregated at the establishment-level and allows us to locate firms and choose single-city firms. The firm identifier is identical to the one present in the customs and balance sheet data thus allowing us to localize the importing and economic performance of firms in a given city.
2. **Customs data**: This file is produced by the customs office, and compiles the imported and exported values and quantities for each firm-destination-eight-digit product category combination. This file covers only trade in goods and therefore covers primarily manufacturing firms but also include retail and wholesale sectors as well as the service sector more generally. The data is subject to some declaration thresholds which varies depending on whether the trade is extra or within EU. For imports from outside the EU, reporting is required from each firm and flow if the imported value exceeds 1,000 Euros (this threshold was lifted in 2010 but does not vary over the sample period). For imports within the EU, there exists several threshold regarding import flows. Import flows have to be reported as long as the firm's annual trade value exceeds 100,000 Euros. There are only minor variation in the declaration thresholds of over the periods (see row 3 for imports of Table 2 in **?**).[48]
3. **BRN: Income and balance sheet**: This file is produced by the French statistical institute (INSEE) on the basis of corporate income tax returns for firms with revenue above than 763 K euros. This dataset has been widely used in the trade literature (see e.g. Mayer et al., 2014). This information at the firm-level is taken from the official form is CERFA 2050-9 which is the French equivalent to the IRS Form 1120 (Corporate Income Tax Return).
4. **Local exchange gepographical coverage data**: INSEE and the regulatory agency (ARCEP) produced a file approximating the geographical coverage of each LE in terms

[48] The most significant shift is an increase in the threshold in 2001 (from 40k to 100k euros). If anything, this change should play against us finding an effect.

of census blocks (IRIS in France, standing for "Ilôts regroupés pour l'information statistique").

5. **Local exchange expansion date**: This information was hand-collected from a website documenting the current and past quality of internet connection at the local level. It contains the exact date for each LE of the update to DSL technology. We were able to check with economists at Orange that our dates matched exactly their data for the year within the sample (1999-2007).

C.2. List of variables

1. **Value of import**: This variable is constructed from the custom file matched with the establishment-level employment data. It is equal to the sum of the values of the imported goods over a given year for all firms located in a given city.
2. **Number of flows**: This variable is constructed from the custom file matched with the establishment-level employment data. It is defined as sum of the number of importing flows good over a given year for all firms located in a given city. An importing flow is defined as the finest level of observation available in the customs data which corresponds to a unique firm $\times$ product (HS-6) $\times$ sourcing country combination.
3. **Average value per flow**: This variable is defined as: Average value per flow $= \frac{\text{Value of import}}{\text{Number of flows}}$.
4. **Density**: This variable is defined as: Density $= \frac{\text{Population census 1999}}{\text{City Area}}$. The population variable is taken from the 1999 Census, the last exhaustive census available in France, which also matches the year of the very beginning of the broadband expansion.
5. **Number of fiscal households and average fiscal income**: source Fichier de l'impôt sur le revenu des communes (IRCOM)
6. **Employment shares**: Number of employees per sector, aggregated from DADS establishment data. The sectors are defined based on a coarsened version of the 2-digit NAF rev. 2003. The sectors are defined as follows: primary sector (2-digit NAF codes: 1 to 14), manufacturing (2-digit NAF codes: 15 to 39), utilities (40 and 41), construction (45), auto repair (50), retail (51 and 52), hotel (55), transport (60 to 63), finance (65 to 71, also includes real estate and rentals) business services (72 to 74), personal services (80 to 97). Overwhelmingly public sectors are excluded (2-digit NAF: 64 and 75).
7. **Sales**: Sales are documented in the BRN files at the firm-level. This is not inclusive of VAT.
8. **Value added**: Value-added is documented in the BRN at the firm-level. It is inclusive of taxes on production that are deductible with respect to the corporate income tax.

D. Additional empirical material

D.1. Further robustness checks

Table 4.9: Specification checks for main specification: asinh(value of imports)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\{d=-5\}$	-0.001	-0.001	-0.015	-0.017	-0.019		
	(0.030)	(0.030)	(0.031)	(0.031)	(0.032)		
$\mathbb{1}\{d=-4\}$	-0.013	-0.016	-0.037	-0.038	-0.030		
	(0.031)	(0.031)	(0.033)	(0.033)	(0.033)		
$\mathbb{1}\{d=-3\}$	0.010	-0.002	-0.028	-0.027	-0.034		
	(0.026)	(0.027)	(0.029)	(0.029)	(0.031)		
$\mathbb{1}\{d=-2\}$	0.046**	0.039*	0.023	0.023	0.034		
	(0.022)	(0.023)	(0.024)	(0.024)	(0.026)		
$\mathbb{1}\{d=0\}$	0.065*	0.071**	0.092**	0.090**	0.090**	0.055	0.085**
	(0.035)	(0.035)	(0.036)	(0.036)	(0.037)	(0.038)	(0.039)
$\mathbb{1}\{d=+1\}$	0.106**	0.110**	0.155***	0.150***	0.150***	0.089*	0.142***
	(0.045)	(0.044)	(0.048)	(0.048)	(0.050)	(0.047)	(0.052)
$\mathbb{1}\{d=+2\}$	0.169***	0.185***	0.251***	0.242***	0.254***	0.160***	0.233***
	(0.059)	(0.057)	(0.064)	(0.064)	(0.069)	(0.059)	(0.067)
$\mathbb{1}\{d=+3\}$	0.204***	0.230***	0.318***	0.304***	0.328***	0.201***	0.295***
	(0.074)	(0.073)	(0.087)	(0.088)	(0.094)	(0.076)	(0.090)
$\mathbb{1}\{d=+4\}$	0.248***	0.283***	0.395***	0.375***	0.381***	0.251***	0.368***
	(0.091)	(0.090)	(0.109)	(0.112)	(0.118)	(0.092)	(0.111)
$\mathbb{1}\{d=+5\}$	0.335***	0.379***	0.516***	0.485***	0.513***	0.342***	0.485***
	(0.119)	(0.118)	(0.140)	(0.145)	(0.152)	(0.120)	(0.143)
R2	0.82	0.82	0.82	0.82	0.82	0.82	0.82
Year FE	Yes	No	No	No	No	No	No
Year x Dep. FE	No	Yes	Yes	Yes	No	Yes	Yes
Year x ZE FE	No	No	No	No	Yes	No	No
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Density controls	No	No	Yes	Yes	Yes	No	Yes
Fiscal + Sector + Educ. controls	No	No	No	Yes	No	No	No
Spec.	D	D	D	D	D	SD	SD
N	178818	178818	178723	178656	178721	178818	178723

NOTES : This table presents event-study estimates based on Equations (4.2) (columns 1 to 5) and (4.3) (columns 6 and 7). Robust standard errors clustered at the département level are presented in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimates are plotted in Figure 4.3, see the figure notes for more details on the controls.

Table 4.10: Static panel-fixed effect model

	Ln (Values of Imports)					
	(1)	(2)	(3)	(4)	(5)	(6)
ADSL access	0.044**	0.047**	0.051**	0.050**	0.058**	0.064***
	(0.021)	(0.021)	(0.020)	(0.020)	(0.023)	(0.022)
Year FE	Yes	No	No	No	No	No
Year x Dep. FE	No	Yes	Yes	Yes	No	No
Year x ZE FE	No	No	No	No	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Density controls	No	No	Yes	Yes	No	Yes
Fiscal + Sector + Educ. controls	No	No	No	Yes	No	No
Spec.	S	S	S	S	S	S
Observations	96947	96947	96916	96916	96936	96905

NOTES : This table presents results of a panel regression where the explanatory variable of interest *Broadband Internet access* is the continuous measure of broadband access in city i and year t defined as a time-weighted percentage of area covered in city i as defined in equation (4.1). Robust standard errors clustered at the province (département)-level are presented in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See the figure 4.3's notes for more details on the controls.

Sensitivity with respect to binning pre-treatment periods

This appendix presents results based on the following specification:

$$Y_{it} = \beta_{d_0} \times \mathbb{1}\{t - t_{0i} \leq -4\} + \sum_{\substack{d=-d_0+1 \\ d \neq -1}}^{d_1} \beta_d \times \mathbb{1}\{d + t_{0i} = t\} + \mathbf{x}'_{it}\delta + \alpha_i + \psi_{d(i),t} + \varepsilon_{it} \quad (4.50)$$

t_{0i} refers to the city-specific arrival year of broadband internet. Periods relative to broadband expansion are index by d. In this specification, the coefficients for $d \leq d_0$ are constrained to be equal (they are binned together in the terminology of Schmidheiny and Siegloch (2019)). More specifically, we consider observations with $d \in \{-6, 5\}$ and we set $d_0 = 4$.

Table 4.11: Specification Check: ln(value of imports) , binning $d \in \{-6, -5, -4\}$

	Ln (Values of Imports)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\{d \leq -4\}$	-0.028	-0.030	-0.046	-0.045	-0.050		
	(0.038)	(0.040)	(0.041)	(0.041)	(0.043)		
$\mathbb{1}\{d = -3\}$	-0.020	-0.029	-0.043	-0.042	-0.046		
	(0.026)	(0.027)	(0.028)	(0.028)	(0.029)		
$\mathbb{1}\{d = -2\}$	-0.016	-0.021	-0.029	-0.029	-0.027		
	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)		
$\mathbb{1}\{d = 0\}$	0.008	0.013	0.025	0.024	0.033*	0.017	0.027
	(0.015)	(0.016)	(0.017)	(0.017)	(0.018)	(0.016)	(0.016)
$\mathbb{1}\{d = +1\}$	0.041	0.052**	0.077***	0.076***	0.084***	0.050**	0.069***
	(0.025)	(0.026)	(0.027)	(0.027)	(0.030)	(0.022)	(0.022)
$\mathbb{1}\{d = +2\}$	0.091***	0.111***	0.151***	0.149***	0.167***	0.102***	0.133***
	(0.034)	(0.035)	(0.039)	(0.040)	(0.044)	(0.028)	(0.030)
$\mathbb{1}\{d = +3\}$	0.106**	0.137***	0.198***	0.195***	0.206***	0.123***	0.170***
	(0.048)	(0.050)	(0.057)	(0.059)	(0.061)	(0.037)	(0.041)
$\mathbb{1}\{d = +4\}$	0.124**	0.171***	0.255***	0.250***	0.260***	0.152***	0.218***
	(0.061)	(0.063)	(0.072)	(0.074)	(0.076)	(0.043)	(0.050)
$\mathbb{1}\{d = +5\}$	0.133*	0.189**	0.301***	0.294***	0.302***	0.164***	0.255***
	(0.075)	(0.077)	(0.089)	(0.092)	(0.094)	(0.055)	(0.064)
R2	0.84	0.85	0.85	0.85	0.85	0.85	0.85
Year FE	Yes	No	No	No	No	No	No
Year x Dep. FE	No	Yes	Yes	Yes	No	Yes	Yes
Year x ZE FE	No	No	No	No	Yes	No	No
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Density controls	No	No	Yes	Yes	Yes	No	Yes
Fiscal + Sector + Educ. controls	No	No	No	Yes	No	No	No
Spec.	D	D	D	D	D	SD	SD
N	96947	96947	96916	96916	96905	96947	96916

NOTES : This table presents event-study estimates based on Equations (4.50)) (columns 1 to 5) and (4.3) (columns 6 and 7). Observations up to 6 year before and 5 year after broadband expansion are included in the estimating sample and the indicator variables for periods -6 to -4 are binned together. Robust standard errors clustered at the département level are presented in brackets. * p < 0.1, ** p < 0.05, *** p < 0.01.

D.2. Explaining internet coverage : table

Table 4.12: Explaining variation in internet coverage: full panel regressions

	(1) Covariates	(2) Twoway FE	(3) (2)+density	(4) (2)+indus.	(5) (2)+trade.	(6) (2)+ all cova.
Lagged % primary	0.250***			-0.00313		0.00963
	(30.15)			(-0.13)		(0.40)
Lagged % construction	0.0377***			0.00890		0.0174
	(9.11)			(0.70)		(1.46)
Lagged % auto	0.0190			0.00536		0.00207
	(1.84)			(0.19)		(0.08)
Lagged % retail	0.0215***			0.00570		0.00715
	(4.42)			(0.39)		(0.54)
Lagged % hotel	0.0154*			-0.00102		-0.00445
	(2.20)			(-0.05)		(-0.21)
Lagged % transport	0.0266***			0.00837		0.0101
	(4.78)			(0.46)		(0.56)
Lagged % finance	0.0365*			0.00804		0.0300
	(2.31)			(0.23)		(0.90)
Lagged % service_pro	0.101***			-0.00520		-0.00459
	(12.98)			(-0.34)		(-0.31)
Lagged % service_pers	0.0188***			0.0390*		0.0562***
	(4.18)			(2.61)		(3.86)
Lagged % utilities	0.0802***			0.0201		0.0476
	(4.05)			(0.34)		(0.83)
Lagged Δ % primary	-0.261***			-0.0509**		-0.00222
	(-17.81)			(-3.10)		(-0.14)
Lagged Δ % construction	-0.00856			-0.00541		-0.0103
	(-0.77)			(-0.58)		(-1.19)
Lagged Δ % auto	-0.0227			-0.0195		-0.0196
	(-0.88)			(-0.98)		(-0.98)
Lagged Δ % retail	-0.000407			-0.000326		-0.00274
	(-0.03)			(-0.03)		(-0.30)
Lagged Δ % hotel	-0.00420			-0.0119		-0.00880
	(-0.24)			(-0.85)		(-0.63)
Lagged Δ % transport	-0.0236			-0.0134		-0.0118
	(-1.58)			(-1.12)		(-1.09)
Lagged Δ % finance	0.0310			0.0278		0.0207
	(1.28)			(1.29)		(1.01)
Lagged Δ % service_pro	-0.0412**			-0.0135		-0.00428
	(-2.69)			(-1.01)		(-0.36)
Lagged Δ % service_pers	0.00304			-0.0190		-0.0292*
	(0.22)			(-1.63)		(-2.54)
Lagged Δ % utilities	-0.0733			0.00602		0.0320
	(-1.45)			(0.15)		(0.79)
Lagged # flows of imports	0.00640***				-0.00810***	-0.00375
	(6.96)				(-3.43)	(-1.70)
Lagged Δ # flows of imports	-0.00265				0.00217	0.000989
	(-1.56)				(1.43)	(0.71)
Lagged value of imports	0.000818**				0.00117*	0.000398
	(2.97)				(2.11)	(0.75)
Lagged Δ value of imports	-0.000199				-0.0000951	0.000177
	(-0.54)				(-0.30)	(0.58)
$\mathbb{1}$ { year=1999 } × Ln Density 1990	-0.0475***					
	(-51.73)					
$\mathbb{1}$ { year=2000 } × Ln Density 1990	-0.0331***		0.0337***			0.0338***
	(-35.96)		(6.93)			(6.94)
$\mathbb{1}$ { year=2001 } × Ln Density 1990	0.00694***		0.109***			0.110***
	(7.51)		(26.23)			(26.37)
$\mathbb{1}$ { year=2002 } × Ln Density 1990	0.0449***		0.162***			0.162***
	(48.70)		(34.28)			(34.25)
$\mathbb{1}$ { year=2003 } × Ln Density 1990	0.0807***		0.170***			0.170***
	(87.53)		(27.10)			(27.03)
$\mathbb{1}$ { year=2004 } × Ln Density 1990	0.121***		0.146***			0.146***
	(131.90)		(18.07)			(17.95)
$\mathbb{1}$ { year=2005 } × Ln Density 1990	0.162***		0.0904***			0.0904***
	(175.98)		(12.36)			(12.26)
$\mathbb{1}$ { year=2006 } × Ln Density 1990	0.188***		0.0374***			0.0374***
	(202.10)		(8.11)			(8.10)
$\mathbb{1}$ { year=2007 } × Ln Density 1990	0.198***		0.0162***			0.0164***
	(213.11)		(5.76)			(5.78)
R^2	0.557	0.786	0.812	0.787	0.786	0.812
Trade and/or industry: F-stat	59.56			2.57	4.78	2.5
Density: F-stat	20076.81		221.55			223.78

NOTES : This table presents panel regression estimates based on equation (4.4). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Fig. 4.16. Number of cities by cohort

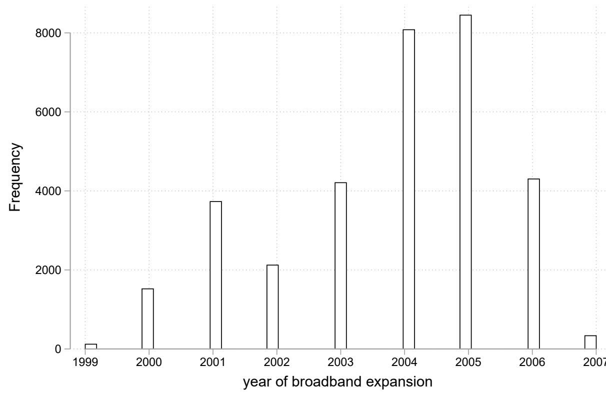
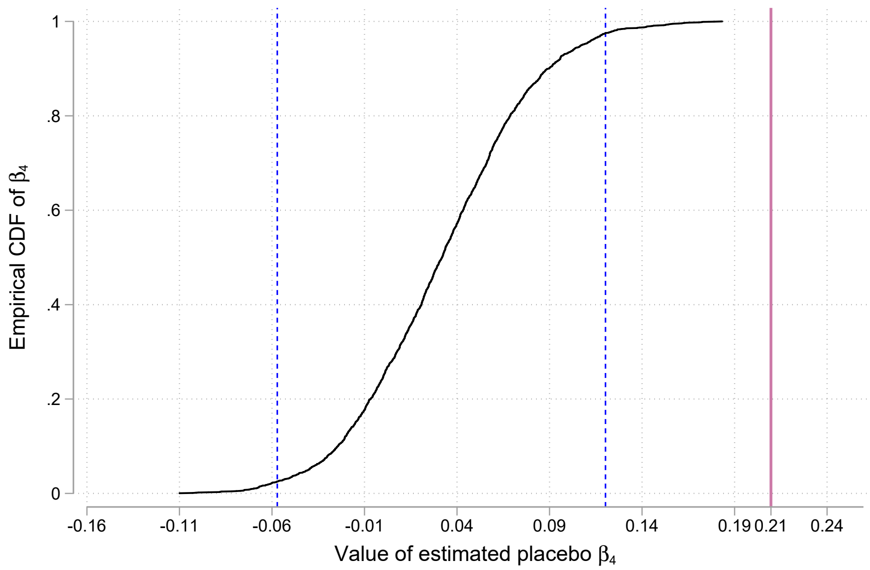
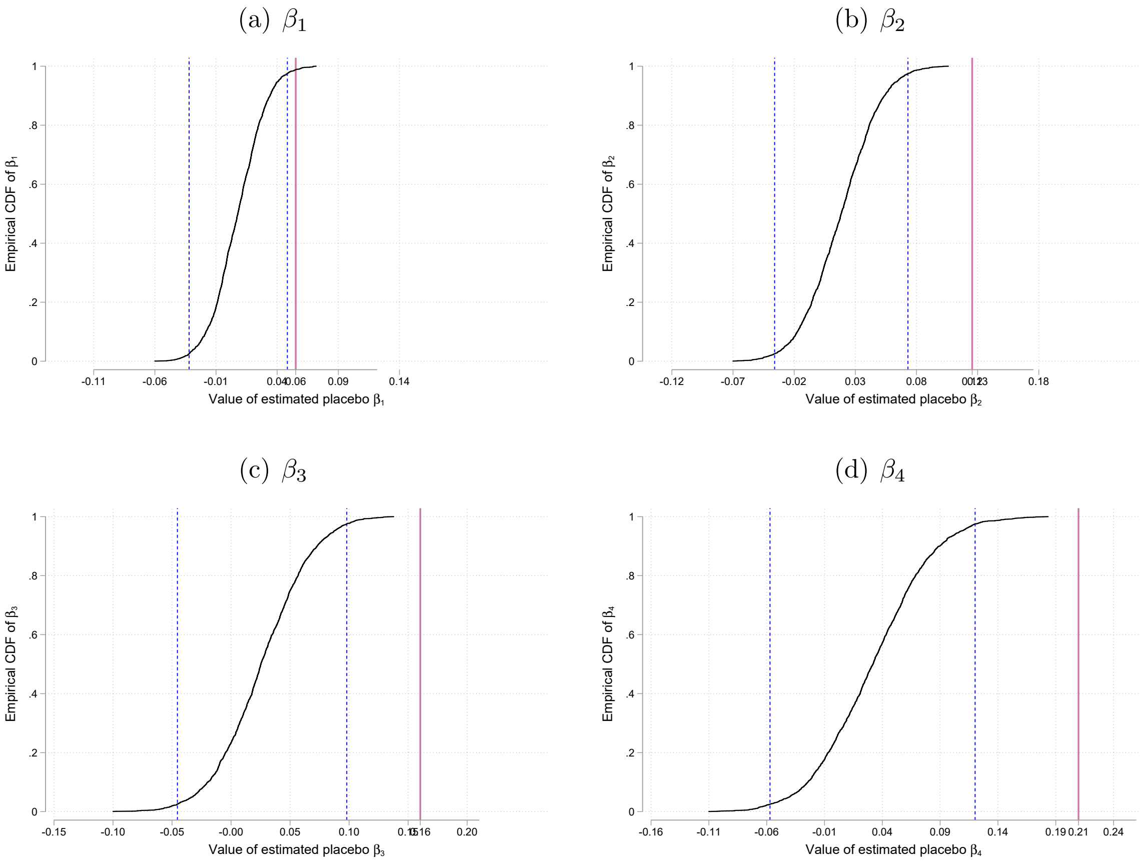


Fig. 4.17. Distribution of Placebo Estimates: Log Imports, β_4



NOTES: This figure plots the empirical cumulative distribution function of placebo estimated effects broadband on log imports, where date of broadband expansion is randomly reallocated across cities within a département (unit of cluster). Draws are with replacement and may include the correct date of treatment. The CDF is constructed from 1000 estimates of β_4 using the specification in equation (4.2) without observable controls. The solid line (in red) corresponds to the actual estimate of the matching specification. It lies outside of the 95% confidence interval that is delineated by the dash lines (in blue).

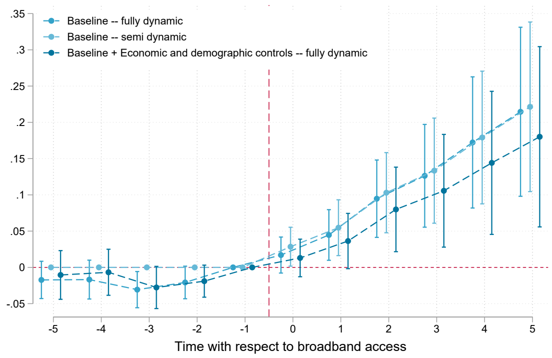
Fig. 4.18. Distribution of Placebo Estimates: Log Imports



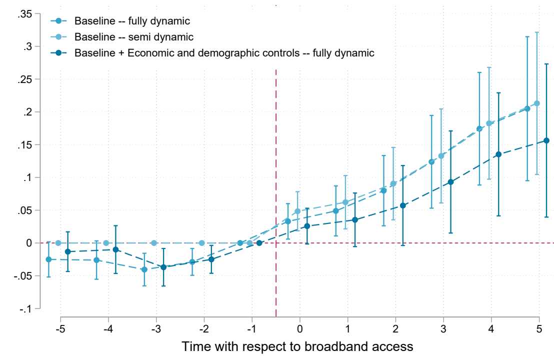
NOTES: This figure plots the empirical cumulative distribution function of placebo estimated effects broadband on log imports, where date of broadband expansion is randomly reallocated across cities within a département (unit of cluster). Draws are with replacement and may include the correct date of treatment. The CDF is constructed from 1000 estimates of β_4 using the specification in equation (4.2) without observable controls. The solid line (in red) corresponds to the actual estimate of the matching specification. It lies outside of the 95% confidence interval that is delineated by the dash lines (in blue).

Fig. 4.19. Including multi-city firms in the analysis

(a) Trade allocated based on HQ location

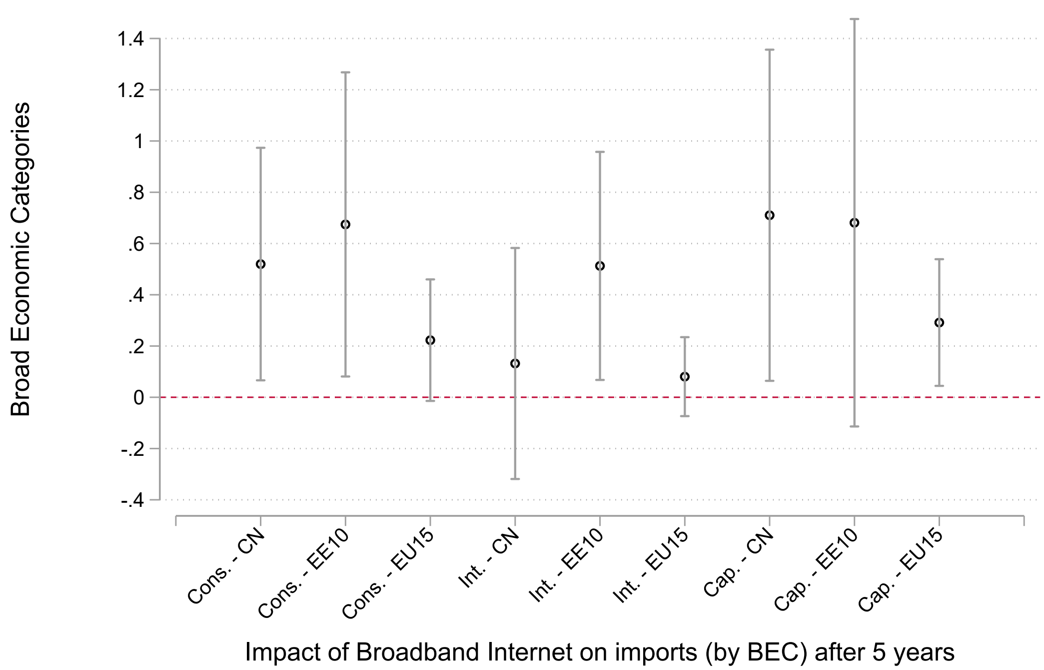


(b) – based on employment



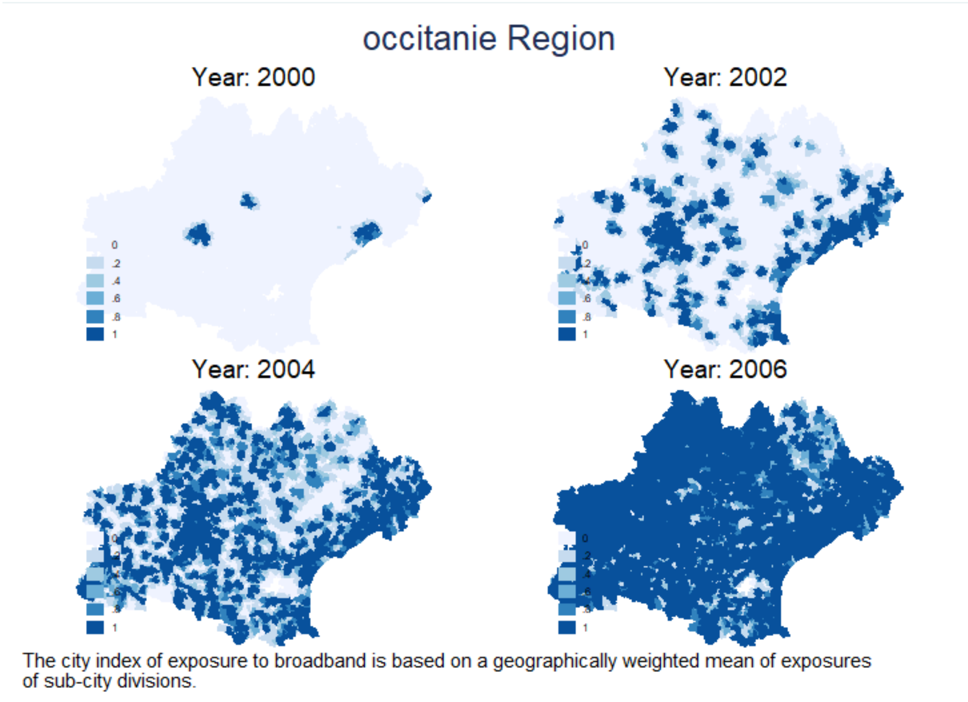
NOTES: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The number of flows (left panel) is defined as the number of the sum of all firm-origin-product combination for a given city in a given year. The average value of per flow (right panel) is defined as the overall value of imports by firms located in a given city divided by the number of flows as defined above. The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

Fig. 4.20. Effect on imports by type of goods and origin countries



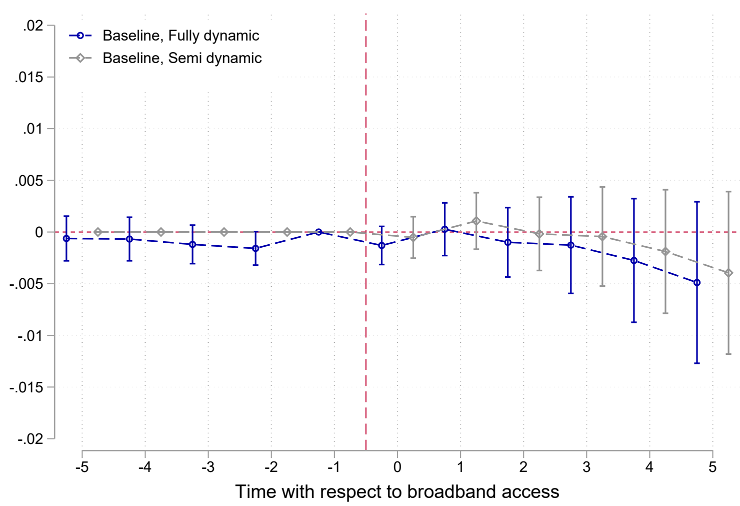
NOTES: This figure plots estimates ($t \geq 0$) for specification in equation (4.3—semi dynamic) and for trade flows grouped by BEC category and origin country. Cons = consumption goods. Int = intermediate goods. Cap. = capital goods. Density controls. The sample include all cities with a positive trade flow (import).

Fig. 4.21. The progressive roll-out of the DSL technology in Occitanie—$\widetilde{Z}$



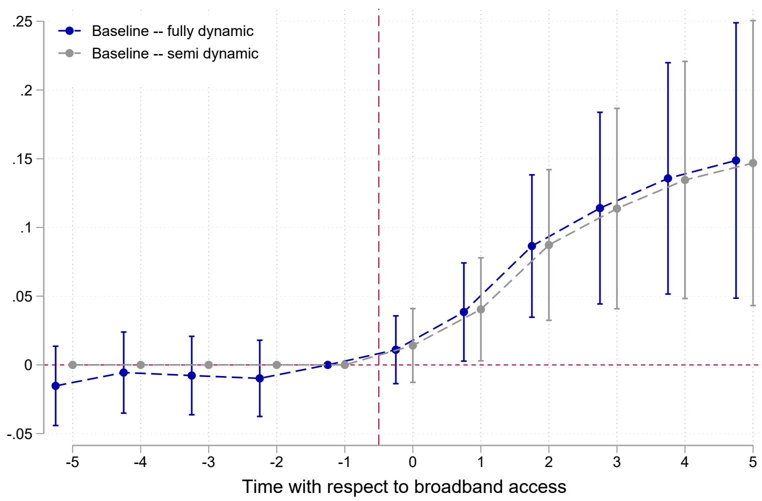
Notes: This figure presents the geographical distribution of the continuous measure of local broadband availability (variable $\widetilde{Z}$) as defined in Equation (4.1).

Fig. 4.22. Intermediate inputs over-sales-ratio



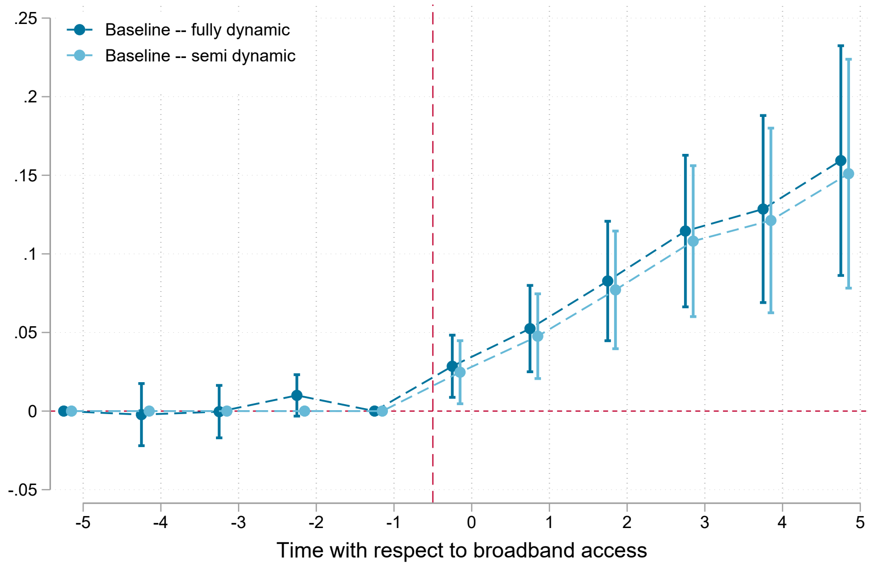
Notes: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The dependent variable is defined as ln(intermediate input / total sales).

Fig. 4.23. Share of foreign inputs (ln), excluding capital goods



Notes: This figure plots estimates for specification in equation (4.2—fully dynamic) and (4.3—semi dynamic). The baseline specification includes 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the département level. The sample include all cities with a positive trade flow (import).

Fig. 4.24. Log of the value of imports at firm-level

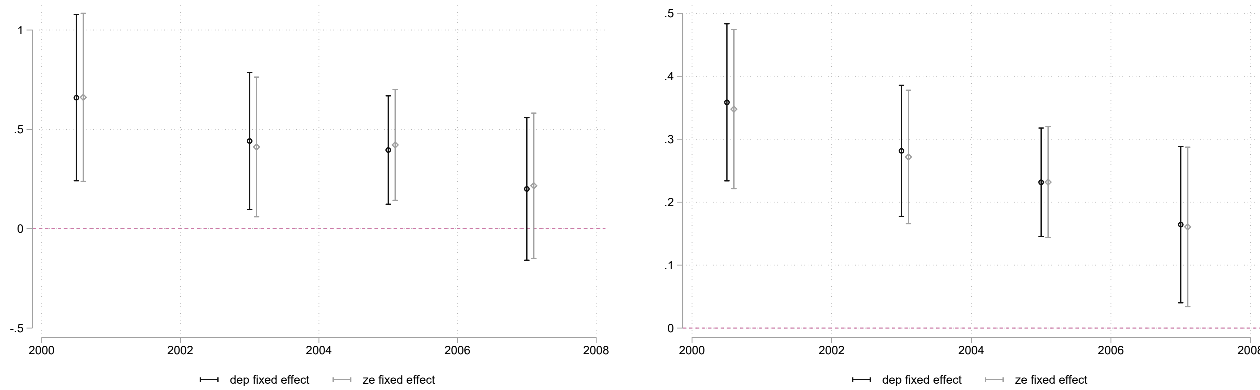


NOTES: This figure plots estimates for firm-level fully dynamic and semi dynamic event study. The sample include all single-city firms with at least one positive trade flow (import) before Broadband Internet access. The baseline specification includes firm, year × département and city FE as well as 1999 population density at the city level interacted with quadratic and linear trends. 95 % confidence interval are presented. Standard errors clustered at the province (département) level.

Fig. 4.25. Long-difference results:

(a) Δ asinh (value of imports)

(b) D-H Growth Rate: $\frac{\Delta \text{M}}{2(M_{1999}+M_{2007})}$



NOTES: The figure reports the estimation of the following equation: $\Delta Y_{i,99,07} = \sum_{c=1999}^{2007} \beta_c \times \mathbb{1}\{i \in c\} + X_i'\delta + \gamma_{r(i)} + \varepsilon_i$ where c refers to the cohort of a city and $\Delta Y_{i,99,07}$ represents a change in a given outcome Y between 1999 and 2007. The parameters $\gamma_{r(i)}$ is a fixed effect taken at the département or commuting-zone (CZ) level. On the left side, the outcome is the change in the asinh() of the value imports, on the right panel it is the Davis-Haltiwanger growth rate, defined as: $\Delta Y_{i,99,07} = (1/2)(Y_{i,07} + Y_{i,99})^{-1}(Y_{i,07} - Y_{i,99})$. Cities are grouped into 4 cohorts corresponding to years {1999,2000,2001}, {2002,2003}, {2004,2005}, {2006,2007}. Two sets of estimates corresponding to different sets of fixed effects are presented : département fixed effects, ZE fixed effects where ZE stands for commuting zone. Confidence interval at the 95% level are displayed around the point estimates, based on robust standard errors clustered at the département-level. Controls include the log of density as of 1990.

E. Conceptual framework: additional material

E.1. Other extensions

Endogenous domestic intermediate. Our quantification takes the price of domestic intermediate inputs has fixed. In order to aggregate the effect of changes in the import environment onto the overall price level in a general equilibrium setting, we need to specify how the intermediate goods sector operates. A natural way to do this is to postulate a roundabout production whereby firms use a domestic input that is produced using the output of all other firms in the economy (Caliendo and Parro, 2015, is a recent example using this technology). In this one-sector model this implies that $P = p_{X_D}$. Accordingly, substituting equation (4.26) into (4.17), we have:

$$P = \frac{\sigma-1}{\sigma}\left(\int_{i\in\Omega}\left(\tilde{A}_i^{-1} sd_i^{\frac{1-\gamma}{\varepsilon-1}}\right)^{1-\sigma} di\right)^{\frac{1}{1-\sigma}} (p_{X_D})^{1-\gamma}\, w^{\gamma}$$

$$\Rightarrow P = \frac{\sigma-1}{\sigma}\left(\int_{i\in\Omega}\left(\tilde{A}_i^{-1} sd_i^{\frac{1-\gamma}{\varepsilon-1}}\right)^{1-\sigma} di\right)^{\frac{1}{(1-\sigma)\gamma}} w$$

Therefore the aggregate effect is now equal to $1/\gamma$ times the micro effect, but this does not affect the relative important of the import-channel. With $\gamma = 0.53$ for instance, we have a total effect of 3.5% and an import channel of 1.41%.

E.2. Conceptual framework: a simple example

In firm-level importing models, the choice of the optimal set of sourcing problem is complex as, unlike for exporting problems, the choice of individual sourcing countries are interdependent as they interact through the cost function (Antras et al., 2017). The approach developed in Section B, inspired by Blaum et al. (2018), allows us to quantify the consumer welfare effect of broadband internet in a firm-level importing model without having to fully characterize the importing environment and solve for the optimal sourcing strategy.

In this appendix section, we illustrate the sufficiency result presented in B.3, within a very simple but fully specified model. When not otherwise specified, the notation follows that of Section B.

Production function follows a Cobb-Douglas:

$$Y = A \times L^{\gamma} X^{1-\gamma}$$

The cost minimization program writes as:

$$\min wL + PX \text{ s.t. } Y = 1$$

The resulting unit cost is:

$$u = A^{-1} w^{\gamma} P^{1-\gamma} \times \gamma^{-\gamma} (1-\gamma)^{-(1-\gamma)}$$

Here X is aggregation from several types of goods in a CES fashion:

$$X = \left(\sum_{c \in \Omega} \alpha_c D_c^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where D_c denotes the quantity of input c. The price of input c writes as p_c. Given cost minimization and the CES production function, we have the following results:

$$
\begin{aligned}
&\min \sum_{c\in\Omega} p_c D_c \text{ s.t. } X = 1 \\
\text{FOC: } & p_c = \lambda \alpha_c D_c^{-\frac{1}{\varepsilon}} \\
&\sum_{c\in\Omega} D_c p_c = \lambda \sum_{c\in\Omega} \alpha_c D_c^{\frac{\varepsilon-1}{\varepsilon}} \times X^{\frac{1}{\varepsilon}} = \lambda X \\
&D_c p_c = (\sum_{c\in\Omega} D_c p_c)^{\varepsilon} \alpha_c^{\varepsilon} p_c^{1-\varepsilon} \Leftrightarrow \sum_{c\in\Omega} D_c p_c = \left(\sum_{c\in\Omega} \alpha_c^{\varepsilon} p_c^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} =: P
\end{aligned}
$$

Next, we consider the case of two inputs: one domestic (1) and one foreign (2). In that context, cost share of input 1 writes as:

$$
s_1 := \frac{D_1 p_1}{D_1 p_1 + D_2 p_2} = \frac{p_1^{1-\varepsilon} P^{\varepsilon} \alpha_1^{\varepsilon}}{P} = p_1^{1-\varepsilon} P^{\varepsilon-1} \alpha_1^{\varepsilon}
$$

Now, we consider D_2. The foreign goods is itself a CES aggregation of several national origins following an Armington-type of model:

$$
D_2 = \left(\sum_{v\in\Theta} h_v^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}
$$

Unit cost will depends on variable cost r_v plus fixed cost f per destination v with convex cost in the number of destination $|\Theta|$ ($\lambda \geq 1$):

$$
u_2 = \sum_{v} r_v h_v + |\Theta|^{\lambda} f
$$

Assuming $r_v = r$ v, the variable cost component writes as:

$$
p_2(\Theta) = \left(\sum_{v\in\Theta} r_v^{1-\eta}\right)^{\frac{1}{1-\eta}} = |\Theta|^{\frac{1}{1-\eta}} \times r
$$

Now, the firm chooses Θ (we assume continuum of countries):

$$u_2 = |\Theta|^{\frac{1}{1-\eta}} \times r + |\Theta|^{\lambda} f$$

The first order condition writes as:

$$\begin{aligned}
&\frac{r}{1-\eta}|\Theta|^{\frac{\eta}{1-\eta}} + \lambda|\Theta|^{\lambda-1} f = 0 \\
\Leftrightarrow &|\Theta| = \left(\frac{\lambda(\eta-1)f}{r}\right)^{\frac{1-\eta}{1-\lambda(1-\eta)}}
\end{aligned}$$

We consider the case where $\lambda = 1$. The unit cost of the foreign inputs, inclusive of fixed and variable trade costs, writes as:

$$u_2 = (f^{\frac{1}{\eta}} r^{\frac{\eta-1}{\eta}}) \underbrace{\left((\eta-1)^{\frac{1}{\eta}} + (\eta-1)^{\frac{1-\eta}{\eta}}\right)}_{\equiv B} = \frac{\eta}{\eta-1} \times f^{\frac{1}{\eta}} r^{\frac{\eta-1}{\eta}} (\eta-1)^{\frac{1}{\eta}}$$

Consequently, the unit cost of the intermediate goods writes as:

$$\begin{aligned}
P &= \left(\alpha_1^{\varepsilon} p_1^{1-\varepsilon} + (1-\alpha_1)^{\varepsilon} u_2^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \\
&= \left(\alpha_1^{\varepsilon} p_1^{1-\varepsilon} + (1-\alpha_1)^{\varepsilon} ((f^{\frac{1}{\eta}} r^{\frac{\eta-1}{\eta}})B)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \\
&= \left(\alpha_1^{\varepsilon} p_1^{1-\varepsilon} + (1-\alpha_1)^{\varepsilon} ((f^{\frac{1-\varepsilon}{\eta}} r^{\frac{(1-\varepsilon)(\eta-1)}{\eta}}) B^{1-\varepsilon})\right)^{\frac{1}{1-\varepsilon}}
\end{aligned}$$

We can then derive the fully specified overall unit cost as a function of fixed cost f and variable cost r of importing:

$$u = w^{\gamma} \left(\alpha_1^{\varepsilon} p_1^{1-\varepsilon} + (1-\alpha_1)^{\varepsilon} ((f^{\frac{1-\varepsilon}{\eta}} r^{\frac{(1-\varepsilon)(\eta-1)}{\eta}}) B^{1-\varepsilon})\right)^{\frac{1-\gamma}{1-\varepsilon}} \times \gamma^{-\gamma}(1-\gamma)^{-(1-\gamma)}$$

We can also express unit cost u as a function of s_1 using its relationship with the price index P:

$$u = A^{-1} w^{\gamma} (p_1/\alpha_1)^{\frac{1-\gamma}{\varepsilon-1}} (s_1)^{\frac{1-\gamma}{\varepsilon-1}} \times \gamma^{-\gamma}(1-\gamma)^{-(1-\gamma)}$$

In that setting the sufficiency result states that any effect of a variation in f, r on the unit cost u is operating through s_1. Therefore, the effect of broadband internet on the unit cost through either changes in fixed cost f or variable cost r of imports is fully summarized by

its effect on s_1: $\frac{d \ln s_1}{dBI} \times \frac{1-\gamma}{1-\varepsilon}$.

Conclusion

This dissertation is a collection of four chapters that study frictional credit and goods markets.

In chapter 1, I develop a new theory of firm-bank matching subject to search frictions. I provide a causal evidence of such frictions affect firm-bank matching and the allocation of bank credit, using the staggered roll-out of broadband internet in France as a shock on transaction and search costs. I show that this technology-induced reduction in search frictions triggers an increase by 6% of the share of credit exchanged between interconnected cities. This positive effect varies with the initial level of search frictions: it is higher when two very distant cities are connected. On the contrary, the effect the effect is almost null when two neighboring cities, already economically very closely tied, are interconnected by internet. Leveraging bank branch-level data, I document that Broadband Internet diffusion allows banks to match with new firms located remotely. Connected banks increase their share of credit lent to firms located outside their city by 10%, and their share of remote clients by almost 12%. As a result, the average distance between a bank and its customers increases by 10% in the medium run after broadband internet access. Finally, I plug these estimates into the equation linking search frictions to loan prices. Interpreted within my model, the reduced-form estimates imply that the reduction in search frictions due to the large diffusion of Broadband Internet lowered the cost of debt for small businesses by 4.9% on average. Overall, this chapter highlight the role of transactions and search cost in shaping firm's access to credit. Credit markets with high search frictions make financing by bank credit both difficult, time-consuming and onerous, especially for small businesses. This conclusion calls for a variety of economic policies aiming at to make the process of searching and applying for credit more fluid, efficient and less burdensome, in particular in a period of pandemic marked by the disappearance of face-to-face interactions and the consequent surge of digitalization.

Chapter 2 studies another aspect of frictions affecting the credit markets and hindering the

ability of firms to raise external finance. We indeed provide evidence that a large share of local bank branches in France are specialized in lending to some industries. We also show that this local bank specialization matters for SMEs' access to credit: a small firm enjoys on average a better access to bank credit when the local branch of its relationship bank has gained a better knowledge of the industry to which the firm belongs. Using a very rich dataset on bank branch closures over 2010-2017, we find that closures on average entail a substantial and persistent loss in credit supplied for the local customers which are transferred to another branch. Importantly however, firms which, as a consequence, lose access to loan officers specialized in their industry are more badly hit than others. This effect is not wiped out when considering confounding factors such as the increased distance or the decreased local bank competition that may be associated with the transfer.Our findings have possibly important policy implications. First, competition regulators overseeing cuts in bank branch networks following bank mergers should be aware that local bank specialization exists and matters for credit supply to SMEs. Concerns may arise if simple competition rules, such as the obligation to close one of two branches in a given area, are implemented bluntly. Second, supervisors and bank managers should be aware that closing specialized branches may have differentiated effects on small borrowers, depending on their industry. Many issues remain however open, and left for further research. A first, looming question is of course to understand how and why local bank branches become specialized in funding one specific industry (more than other neighbor branches do). A possible cause of specialization may be diffusion effects within networks of customers (e.g., buyers and suppliers, or members of local professional clubs): in other words, the lending pattern of a branch is likely to be history-dependent. An idiosyncratic role of some loan officers, who e.g. join a branch with a better experience in dealing with a certain type of firms, also comes to mind. In this case, the pending issue for bank managers is to ensure that such human capital is not lost whenever the branch closes.

In chapter 3, we study credit markets (and how frictions shape the aggregate fluctuations of credit) from a macroeconomic standpoint. Our analysis highlights the role and importance of the extensive margin in aggregate credit fluctuations. The methodology we develop for relationship flows extends that of labor research to account for the specificities of credit markets and is applicable to the study of other frictional markets and countries with available credit register data. We view our empirical approach and dataset as a novel laboratory and as a first step toward uncovering more properties of credit relationships and their aggregate implications. While we focus mostly on establishing stylized facts and identifying the distinctive features of extensive/intensive margins, we believe that fleshing out the potential economic mechanisms behind these dynamics can provide additional connections to the macro-finance

literature, and in particular the role played by collateral and bank balance-sheet channels. More broadly, given the key role played by the extensive margin of credit along the business cycle frequency and in the long run, this analysis raises the issue of whether banking models abstracting from such a quantitatively important dimension provide a reasonable benchmark for the study of aggregate credit fluctuations. Thus, building models that account for both margins is, in our opinion, critical when thinking about bank credit. We leave the implications of these arguments for future research.

Finally, chapter 4 studies how a technological shock may affect international goods markets by reducing information frictions. We find broadband expansion to have substantially changed the importing patterns of French small and medium-sized firms. The affected firms increased their overall imports by around a quarter. This was driven by a rise in both the number of source countries and the count of products imported. The rise in imported value exceeded the growth of turnover, resulting in a large increase in import intensity (the imports over sales ratio goes up by about 15%). The first implication of our findings relates to the debate regarding trade vs technology as separate explanations for labor market outcomes. The literature argue that trade shocks had more impact than technological shocks on the decline of US manufacturing employment over the 2000's. Our work suggests that it is hard to empirically disentangle the effects of trade from those of technology: a specific instance of technological change, namely broadband internet, has contributed to the trade shocks that this literature typically takes as given. This interaction is quantitatively large, since we estimate that the rise in import penetration in France over the 1997-2007 period would have been 16% smaller without broadband internet.

Institut d'études politiques de Paris
ÉCOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'économie
Doctorat en sciences économiques

Les frictions informationnelles dans les marchés du crédit et des biens

Clément MAZET-SONILHAC

Thèse dirigée par

Thomas CHANEY, Professeur des Universités, Sciences Po

soutenue le 25 juin, 2021

Jury

Mme Cecilia BUSTAMENTE, Professor of Economics, University of Maryland *(rapportrice)*
Mr Thomas CHANEY, Professeur des Universités, Sciences Po
Mr Thierry MAYER, Professeur des Universités, Sciences Po
Mme Isabelle MÉJEAN, Professor of Economics, Ecole Polytechnique *(rapportrice)*

Résumé

Les frictions informationnelles dans les marchés du crédit et des biens

"One should hardly have to tell academicians that information is a valuable resource: knowledge is power. And yet it occupies a slum dwelling in the town of economics. Mostly it is ignored."

– George J. Stigler, *The Economics of Information*

Le lauréat du Prix Nobel 1982, George J. Stigler, remarquait dans un article fondateur intitulé *The Economics of Information* (Stigler, 1961) l'absence criante de concept d'*information* dans la littérature économique. Ce constat est aujourd'hui obsolète: en soixante ans, la notion d'information a pénétré de nombreux champs de la recherche économique. Nous savons désormais que l'information est imparfaite, qu'il est parfois couteux d'en acquérir et qu'il existe d'importantes asymétries d'information. Les quatre chapitres de cette thèse sont une tentative d'étendre le champs nos connaissances sur les effets de l'information imparfaite dans les marchés du crédit et des biens, en s'appuyant sur des données micro-économiques. Dans cette introduction, je présente d'abord le contexte de chaque chapitre séparément. Une description détaillée des chapitres est ensuite fournie.

Un des premiers courants de la littérature économique étudiant l'idée d'information s'est concentré sur l'*acquisition d'information* en développant des modèles de *search* dans lesquels il est coûteux pour les individus d'acquérir une information privée (McCall, 1970; Diamond, 1971), s'éloignant ainsi des modèles d'équilibre général walrasiens qui reposent sur l'idée d'un ajustement instantané entre l'offre et la demande. Ce champs de recherche a progressivement évolué vers les modèles d'appariement et de négociation (*matching and bargaining*) et de recherche dirigée (*directed search*) qui sont aujourd'hui largement utilisés pour étudier le marché du travail (cf. Rogerson et al., 2005), l'économie monétaire (Kiyotaki and Wright, 1993), l'économie financière (Duffie et al., 2005; Weill, 2007), l'économie urbaine (cf. Zenou, 2009 pour une monographie) ou encore l'analyse du marché matrimonial (Mortensen, 1988; Shimer and Smith, 2000). Une branche émergente de cette littérature a récemment mis en évidence l'importance des frictions de recherche dans le commerce (international) des biens (Rauch, 2001; Chaney, 2014; Allen, 2014; Lenoir et al., 2018), reconnaissant qu'il faut du temps et des ressources à un exportateur pour se renseigner sur les conditions du marché

et trouver des clients à l'étranger ou, symétriquement, à un importateur pour trouver le bon fournisseur. En outre, cette littérature a également montré comment la diffusion des technologies de l'information et de la communication (TIC) peut réduire ces frictions (Jensen, 2007; Aker, 2010; Goyal, 2010; Lendle et al., 2016; Steinwender, 2018; Akerman et al., 2018; Bhuller et al., 2019). Les chapitres 1 and 4 de cette thèse contribuent directement à cette branche de la littérature.

Dans le chapitre 1, j'étends l'étude des frictions de recherche aux marchés du crédit. Motivé par des évidences empiriques que je documente sur les marchés locaux du crédit en France, je propose une théorie de l'appariement entre entreprises et banques soumis à des frictions de recherche. Les entreprises amorçent un processus de recherche coûteux pour localiser et rencontrer le bon partenaire bancaire. Après la rencontre, les *agency frictions* empêchent les banques de sélectionner et de suivre les projets de manière optimale. J'estime structurellement mon modèle sur des données françaises en utilisant le déploiement progressif de l'Internet haut débit (ADSL), de 1997 à 2007, comme un choc qui réduit les frictions de recherche. Je confirme les prédictions du modèle selon lesquelles l'allocation optimale du crédit a été modifiée par ce choc. Enfin, le modèle permet de quantifier l'impact de cette réduction des frictions de recherche – induite par un choc technologique – sur les prix des prêts. Je constate en effet que l'accès à l'Internet à haut débit a réduit le coût de la dette des petites entreprises de 4,9% sur la période. Cette étude est directement liée au chapitre 4, co-écrit avec Clément Malgouyres et Thierry Mayer, dans lequel nous étudions le rôle de l'Internet haut-débit dans la réduction des frictions de recherche auxquelles sont confrontés les importateurs français. Nous documentons la présence de commerce "induit par la technologie" en France entre 1997 et 2007 et évaluons son impact sur le bien-être des consommateurs. Nous utilisons le déploiement progressif de l'Internet à haut débit pour estimer son effet causal sur le comportement d'importation des entreprises concernées. En utilisant un modèle d'*event study*, nous montrons que l'expansion du haut débit augmente les importations au niveau des entreprises d'environ 25 %. Nous constatons également que la marge "sous-extensive" (nombre de produits et de pays d'approvisionnement par entreprise) est le principal canal d'ajustement et que l'effet est plus important pour les biens d'équipement.

Parrallèlement à l'étude de l'*acquisition d'information*, une littérature largement distincte a identifié l'importance cruciale des *asymmétries d'information* (Akerlof, 1970; Spence, 1973; Stiglitz, 1975; Stiglitz and Weiss, 1981) qui affectent les transactions lorsque l'un des agents économiques dispose de plus d'information – ou d'une meilleure information – que l'autre. En particulier, les idées fondamentales sur les asymétries d'information développées dans les années 1970 continuent de jouer un rôle clé en économie bancaire et en finance d'entreprise.

Cette littérature a en effet mis en évidence l'importance des relations de long-terme entre les banques et entreprises dans l'atténuation des *agency frictions* et donc dans l'allocation du crédit (cf. Boot, 2000; Degryse et al., 2009; Udell, 2015). Les banques, en développant des relations étroites avec les emprunteurs au fil du temps, augmentent leurs capacités à suivre et à contrôler contrôler les activités de leurs clients. Dans le chapitre 2, co-écrit avec Anne Duquerroy, Jean-Stéphane Mésonnier et Daniel Paravisini, nous montrons que la différenciation et la spécialisation peuvent également permettre aux banques de réduire les asymétries d'information et ainsi de gagner en pouvoir de marché. En utilisant des données micro-économiques de relations entre les banques et de petites entreprises (PME) en France, nous montrons que les banques se spécialisent localement (au niveau de l'agence bancaire) par industrie, et que cette spécialisation affecte le montant d'équilibre des crédits octroyés aux petites entreprises. Pour identifier l'impact de la spécialisation, nous exploitons la réallocation des clients d'agences bancaires fermées vers des agences voisines (de la même banque), ce qui provoque une variation quasi-aléatoire dans la correspondance entre l'industrie d'une entreprise et l'industrie de spécialisation de l'agence prêteuse. Nous montrons que la réaffectation des agences bancaires entraîne, en moyenne, une baisse substantielle et permanente du crédit octroyé aux petites entreprises. Cette baisse est deux fois plus importante pour les entreprises dont les comptes sont réaffectés vers des agences moins spécialisées dans leur secteur que leur agence d'origine.

Finalement, dans le chapitre 3, co-écrit avec Yasser Boualam, nous étudions l'effet des frictions informationnelles sur l'allocation du crédit aux entreprises, d'un point de vue macro-économique. Alors que la plupart des modèles en macro-finance font abstraction de la nature de long terme des relations de crédit et de toutes les frictions qui empêchent les banques de former ou de rompre sans coût ces relations de crédit, nous proposons une nouvelle perspective macro-économique sur le processus d'intermédiation du crédit. Nous fournissons en particulier des évidences empiriques sur les rôles clés et distincts joués par les marges intensive et extensive dans les fluctuations agrégées du crédit. Nous posons des questions simples, mais de premier ordre telles que : (i) Lorsque le crédit bancaire agrégé diminue de 5%, est-ce parce que la taille moyenne des prêts (c'est-à-dire la marge intensive) diminue de 5%, ou est-ce parce que 5% des relations banque-entreprise (c'est-à-dire la marge extensive) sont détruites ? (ii) L'origine des fluctuations du crédit agrégé a-t-elle une importance ? (iii) Les chocs de politique monétaire ont-ils un impact différent sur ces marges ?

Chapitre 1: Les frictions de recherche dans les marchés du crédit

Il est coûteux pour les entreprises de trouver le bon partenaire bancaire, en particulier pour les petites et moyennes entreprises (PME) qui consacrent du temps et des ressources à ce processus de recherche. Les petites entreprises multiplient en effet les demandes de prêt (2,7 en moyenne) et entreprennent un processus de demande fastidieux : en moyenne, plus de 33 heures sont consacrées aux formalités administratives de demande de prêt. Globalement, environ un tiers des PME déplorent un processus de demande de crédit difficile et long. Alors que l'effet des asymétries d'information sur l'attribution des crédits a été bien documenté (cf. Akerlof, 1970; Stiglitz and Weiss, 1981; Petersen and Rajan, 1995), on sait peu de choses sur la façon dont les frictions de recherche affectent l'appariement banque-entreprise et l'accès au crédit. Comprendre le rôle des frictions de recherche est particulièrement important pour les décideurs politiques, non seulement parce que les récents développements en matière de technologie de l'information et de numérisation dans le secteur bancaire sont susceptibles d'affecter les frictions de recherche, mais aussi parce que les politiques qui réduisent les coûts de recherche peuvent différer sensiblement des politiques qui réduisent les frictions d'agence traditionnelles.

La première contribution de cet article est de développer une théorie de l'appariement (*matching*) entreprise-banque soumis à des frictions informationnelles. Cette théorie est motivée par de nouveaux faits stylisés que je documentes sur les marchés locaux du crédit en France : j'utilise pour cela une base de données unique sur les prêts aux petites entreprises de 1998 à 2005 qui me permet de mettre en exergue plusieurs faits stylisées sur la dispersion des taux des prêts, l'hétérogénéité des agences bancaires et les flux de crédit entre villes qui suggèrent, en plus de données d'enquête, la présence de frictions de recherche sur les marchés du crédit. Je développe ensuite un modèle d'équilibre partiel d'appariement entreprise-banque qui explique les faits observés dans les données.[1] Le modèle comporte une double hétérogénéité (agences bancaires et entreprises) et des frictions d'information de deux types. Premièrement, les frictions *de recherche* entravent la capacité des entreprises à localiser et à trouver le bon partenaire financier. Deuxièmement, lors de l'appariement, les *asymétries d'information* affectent la capacité des banques à sélectionner et à monitorer correctement les projets. En ajoutant cette structure au processus de recherche et d'appariement entre firmes et banques, je génère un certain nombre de prédictions théoriques liant le niveau des frictions de

[1] D'un point de vue de la modélisation, mon approche est similaire à Eaton et al. (2018) et Lenoir et al. (2018).

recherche (i) au coût de la dette pour les petites entreprises, (ii) aux flux de crédit entre les villes et (iii) à la dynamique de l'appariement entreprise-banque. En particulier, l'intuition du modèle peut se résumer ainsi: lorsque les coûts de recherche diminuent, les entreprises rencontrent davantage de prêteurs potentiels et finissent par emprunter à un taux plus faible.

La deuxième contribution de l'article consiste à confronter le modèle aux données micro économiques. En pratique, je teste les principales prédictions de mon modèle en utilisant la diffusion progressive de l'Internet haut-débit (BI) en France, de 1999 à 2007, comme un choc sur les frictions de recherche. J'utilise ensuite ces résultats d'estimations empiriques pour quantifier une réduction du coût de l'endettement des PME françaises induite par la technologie. La large diffusion des technologies de l'information et de la communication a en effet représenté un changement profond pour le secteur bancaire et l'Internet à haut débit a été le catalyseur de cette transformation numérique. Avec la numérisation, les coûts de transaction ont diminué et l'essor des services financiers en ligne a permis aux entreprises de rechercher le meilleur partenaire bancaire de manière plus rapide et plus efficace, ce qui a entraîné des changements structurels sur le marché bancaire (Hauswald and Marquez, 2003). Kroszner and Strahan (1999) a montré comment l'adoption massive des technologies de l'information a réduit la dépendance de la proximité géographique entre les clients et les banques, et Petersen and Rajan (2002) a documenté l'érosion de la nature locale des prêts aux petites entreprises, avec une distance croissante entre celles-ci et leurs prêteurs aux États-Unis, mais aussi avec de nouvelles habitudes de communication. Des tendances similaires sont observées en France : les flux de crédit inter régionaux ont augmenté de 15% et la distance moyenne entre les entreprises et les banques a augmenté de 10% entre 1998 et 2005. Mon approche empirique fournit une interprétation causale de ces faits, suggérant que les innovations dans les technologies de l'information - la diffusion de l'Internet à haut débit - ont réduit le rôle des coûts de transaction et de recherche dans l'allocation du crédit, permettant aux entreprises de rechercher des crédits plus loin et conduisant à des changements structurels dans les marchés de crédit locaux.

L'identification de l'effet causal de l'adoption d'une nouvelle technologie sur l'appariement entreprise-banque et l'allocation du crédit est difficile en raison de son endogénéité. Les données et le contexte français me permettent de progresser dans cette direction. La troisième contribution de cet article est ainsi de proposer une stratégie d'identification innovante. Celle-ci repose sur une variable instrumentale pour le *timing* de l'expansion de l'Internet à haut-débit, basée sur un plan théorique d'investissement dans les infrastructures de télécommunication. J'utilise des données sur la disponibilité de l'Internet haut-débit au niveau des municipalités sur la période 1999-2007 compilé par Malgouyres (2017) et je le com-

bine avec des informations concernant la densité de population et les infrastructures de télécommunications existantes, à savoir les *boucles locales* en cuivre et les câbles à fibre optique. Comme les opérateurs de l'Internet haut débit s'appuient sur ces infrastructures déjà existantes pour déployer progressivement le haut débit, je montre que le *timing* optimal de l'arrivée du haut débit peut être prédit à l'aide de ces données, sans informations supplémentaires concernant les conditions économiques locales.

Mes résultats montrent de manière causale que les frictions de recherche affectent l'appariement entreprise-banque et l'allocation du crédit bancaire. Premièrement, je constate que les flux de crédit entre les villes suivent une équation de gravité qui est déformée par le déploiement progressif de l'ADSL. La réduction des frictions de recherche induite par ce choc technologique provoque en effet une augmentation moyenne de 6% de la part du crédit échangé entre les villes inter connectées. Conformément aux prédictions du modèle, cet effet varie considérablement avec le niveau initial des frictions de recherche : il est plus élevé lorsque deux villes très éloignées sont connectées. Au contraire, l'effet est négatif lorsque deux villes voisines, déjà très liées économiquement, sont connectées par internet. En termes de robustesse, je mène un exercice de simulation pour tester la performance de l'estimateur Poisson pseudo-maximum likelihood (PPML) utilisé pour l'estimation des équations de gravité dans un cadre dynamique (panel) et avec de nombreux zéros, étendant ainsi le travail de Santos Silva and Tenreyro (2006, 2011). Je confirme ainsi que l'estimateur PPML avec effets fixes n'est pas biaisé et estime de manière satisfaisante l'effet d'un choc variant dans le temps, même avec plus de 90 % de zéros. Je montre également que mes résultats de base ne sont pas affectés par l'ajout d'effets fixes *dyadiques* et ni par l'ajout de la variable dépendante retardée comme covariable.

Mes principaux résultats reposent sur des régressions effectuées au niveau des unités urbaines. Cependant, certaines prédictions du modèle à un niveau moins agrégé nécessitent l'utilisation de données au niveau des agences bancaires. En testant ces prédictions à un niveau désagrégé, je montre que la diffusion de l'Internet à haut débit permet aux banques de financer de nouvelles entreprises situées dans des marchés éloignés. Les banques connectées augmentent de 10% leur part de crédit aux entreprises situées en dehors de leur ville, et de près de 12% leur part de clients éloignés. Par conséquent, la distance moyenne entre une banque et ses clients augmente de 10%, 5 ans après l'accès à l'Internet haut débit. Ces résultats sont robustes à plusieurs menaces potentielles d'identification. Mes résultats d'estimation dynamiques au niveau des agences bancaires ne révèlent aucune tendance (*pre-trend*) avant la date de connexion à l'ADSL et l'ajout de variables de contrôle n'affecte pas mes estimations. Enfin, je quantifie les implications de mes résultats empiriques sur le coût de la dette pour les

petites entreprises, à travers le prisme de mon modèle. En pratique, j'intègre mes estimations empiriques dans l'équation reliant les frictions de recherche aux prix des prêts. Interprétées dans le cadre de mon modèle, les résultats de mes estimations impliquent que la réduction des frictions de recherche due à la diffusion de l'ADSL a fait baisser le coût de la dette des petites entreprises de 4,9% en moyenne. Cette réduction du coût de la dette présente une hétérogénéité spatiale intéressante. Elle est plus forte dans les zones rurales et les villes moyennes que dans les grandes villes françaises. Les entreprises initialement situées loin des agences bancaires, ou qui ne disposaient pas d'une grande variété de contacts bancaires potentiels, bénéficient donc davantage de la réduction des frictions de recherche, car elle leur permet de trouver de nouveaux ou de meilleurs partenaires bancaires. A cet égard, ce chapitre suggère que la diffusion de l'Internet haut-débit réduit les inégalités spatiales d'accès au crédit.

Chapter 2: La spécialisation locale des banques

Dans ce chapitre, co-écrit avec Anne Duquerroy, Jean-Stéphane Mésonnier et Daniel Paravisini, nous montrons comment la différenciation et la spécialisation bancaire permettent aux banques de réduire les asymétries d'information et d'augmenter leur pouvoir de marché.

Les fermetures et consolidations d'agences bancaires en Europe et aux États-Unis après la Grande Récession ont relancé le débat politique et universitaire sur la nature et les implications du pouvoir de marché des banques. Un vaste corpus de travaux théoriques et empiriques, commençant avec Rajan (1992), a exploré une des sources de ce pouvoir de marché : le monopole informationnel obtenu grâce aux prêts dits relationnels (*relationship lending*). Moins étudié – mais d'une importance potentiellement égale – le pouvoir de marché acquis par la différenciation et la spécialisation est le sujet de ce chapitre. L'idée est la suivante: les prêteurs, qui semblent se livrer une concurrence féroce sur un marché du crédit, pourraient en fait jouir d'un grand pouvoir de marché dans certains segments du marché en adaptant leurs produits et services à des clients, des industries ou des types de financement particuliers. Une telle segmentation du marché du crédit peut avoir des implications de premier ordre sur l'accès au crédit des petites entreprises opaques et dépendantes des banques, ainsi que sur son coût. Documenter dans quelle mesure les banques se spécialisent sur un marché segmenté du crédit aux petites entreprises, et évaluer si la spécialisation confère un pouvoir de marché, est un problème de taille en matière de données et mais également d'identification, que nous tentons de résoudre dans ce chapitre.

Nous utilisons des données réglementaires uniques qui contiennent, pour l'univers des re-

lations banques-entreprises en France, l'identité et la localisation de la succursale bancaire fournissant le crédit. Grâce à ces données, nous construisons des mesures de la spécialisation sectorielle des banques et des succursales bancaires sur des marchés géographiques du crédit. La figure 2.1 fournit des éléments de motivation pour notre étude. Elle représente deux mesures différentes de la concentration du marché du crédit, par unité urbaine. La figure de gauche montre la mesure de concentration standard, calculée à partir de la part totale des prêts. La figure de droite montre la moyenne des concentrations de crédit calculées par secteur d'activité, qui tient compte de la segmentation du marché. La différence entre ces deux mesures est théoriquement plus importante lorsque le marché du crédit est segmenté par industrie (par exemple, les deux mesures seront identiques si tous les portefeuilles de prêts bancaires ont la même composition par industrie). En effet, la fraction des unités urbaines présentant un niveau de concentration très élevé (HHI > 0,4) passe de 21% lorsqu'elle est mesurée de manière traditionnelle, à 49% lorsqu'elle est mesurée en tenant compte de la segmentation du crédit par industrie (cf. tableau 2.9 pour plus de détails). Ce fait stylisé indique une forte segmentation sectorielle sur le marché du crédit bancaire aux petites entreprises.

L'objectif principal de notre analyse empirique est d'explorer les implications de la segmentation du marché du crédit sur l'accès des petites entreprises au crédit. L'hypothèse sur laquelle repose cette idée est que, pour une entreprise, l'élasticité de substitution du crédit provenant de banques différentes est plus faible lorsque les banques spécialisées offrent des services différenciés. Par exemple, une entreprise du secteur de la construction trouvera plus difficile ou plus coûteux de substituer le crédit obtenu auprès de la banque (ou de la succursale) spécialisée dans le secteur de la construction que le crédit obtenu auprès d'une banque (ou d'une succursale) généraliste. Une première étape nécessaire consiste à évaluer l'unité d'analyse pertinente pour étudier la spécialisation. En d'autres termes, les banques se spécialisent-elles entièrement par secteur d'activité ? Ou bien la spécialisation est-elle un phénomène local, au niveau de l'agence ?

Pour répondre à cette question, nous adoptons l'approche développée dans Paravisini et al. (2017) pour identifier le secteur de spécialisation des banques, en repérant des parts de portefeuille dédiées à une industrie particulière anormalement élevées. L'intuition de la mesure nécessite un exemple, par soucis de clarté. Supposons que 20% des crédits bancaires d'une zone urbaine soient destinés au secteur de la construction et soient accordés par cinq banques. Les banques sont hétérogènes : alors que quatre banques allouent moins de 10% de leur portefeuille de prêts à la construction, la cinquième banque alloue plus de 40% de son portefeuille de crédits à ce secteur. Cette cinquième banque sera identifiée comme

spécialiste du secteur de la construction pour cette zone urbaine. L'avantage d'utiliser les parts de portefeuille pour détecter les banques spécialisées est que l'identification du secteur de spécialisation n'est pas affectée par la taille du secteur ou par la part de marché de chaque banque dans un endroit donné.

Deux faits stylisés importants ressortent de cet exercice : les succursales bancaires ont tendance à se spécialiser par secteur, mais les différentes succursales d'une même banque présentent généralement des spécialisations industrielles différentes. Plus d'un tiers des agences bancaires en France sont spécialisées dans l'octroi de crédits aux petites entreprises dans au moins un secteur d'activité spécifique. La plupart des zones urbaines comptent des agences bancaires spécialisées. De plus, nous observons que la plupart des industries sont représentées par des agences bancaires spécialisées au niveau local. Par exemple, environ 9% des agences bancaires présentes dans notre échantillon en 2017 sont spécialisées dans le financement des activités de transport et de stockage. Globalement, cela implique qu'une PME française a une probabilité non négligeable d'être connectée à une agence spécialisée dans son type d'activité. Lorsque nous étudions les schémas de spécialisation des agences au sein des banques, nous constatons que les grandes banques se caractérisent par une part importante de succursales spécialisées (37% pour la banque moyenne comptant plus de 10 succursales). Cependant, au sein d'une même banque, les différentes succursales ont tendance à être spécialisées dans des secteurs différents. En bref, la spécialisation sectorielle apparaît dans les données comme un phénomène répandu mais local, qui intervient au niveau des agences bancaires.

Motivés par ces faits stylisés nouveaux, nous nous tournons vers la mesure de l'hétérogénéité de l'élasticité de substitution du crédit des entreprises en fonction de la spécialisation des branches. Notre cadre empirique exploite les réaffectations des emprunteurs entre agences bancaires en raison des fermetures de ces agences. Parmi les agences bancaires actives dans le domaine des prêts aux PME, quelque 700 agences ont été fermées au cours de notre période d'étude (entre 2010 et 2017) dans l'ensemble du pays, en raison de plans de restructuration interne des activités de détail des principales banques. Les fermetures d'agences n'ont pas mis fin aux relations entre les banques et les emprunteurs : tous les comptes de prêts d'une agence en cours de fermeture ont été transférés vers des succursales voisines de la même banque. La réaffectation des agences induit ainsi une variation de la correspondance entre le secteur d'activité de l'emprunteur et le secteur de spécialisation de la succursale, que nous exploitons pour mesurer l'hétérogénéité de l'élasticité de substitution du crédit. Dans l'exemple de l'entreprise de construction ci-dessus, lorsque les services de l'agence sont segmentés par secteur, le transfert du compte de l'entreprise vers une autre

agence généraliste devrait réduire le montant d'équilibre du crédit utilisé par l'entreprise, par rapport à un contrefactuel dans lequel le compte serait transféré vers une autre succursale également spécialisée dans le secteur de la construction. Les fermetures de succursales se sont produites par grandes vagues, et l'identité des succursales fermées et absorbées a été sélectionnée par le siège en fonction de critères tels que la densité bancaire locale, sans doute sans rapport avec la demande de crédit des entreprises individuelles. La nature très désagrégée des données permet également d'utiliser des spécifications dites saturées en effets fixes pour absorber les chocs locaux (au niveau de l'unité urbaine), les chocs au niveau des banques et les chocs au niveau des entreprises qui pourraient se produire en même temps que la fermeture de la succursale.

Nous mettons d'abord en exergue une baisse significative du total des crédits accordés par une banque à une petite entreprise lorsque le compte de cette dernière est réaffecté à une nouvelle agence. En incluant les lignes de crédit non utilisées, le crédit total diminue de 12% en moyenne au cours des trois années suivant la fermeture effective. Une partie de cette baisse est remplacée par une augmentation des crédits accordés par d'autres banques. Cependant, le crédit total de l'entreprise moyenne diminue de façon permanente d'environ 4% après un transfert de compte, par rapport aux autres entreprises du même marché géographique étroit et du même secteur. Nous documentons ensuite l'hétérogénéité de cette baisse du crédit d'équilibre en fonction de la correspondance entre le secteur d'activité de l'emprunteur et le secteur de spécialisation des agences qui ferment et absorbent. Nous constatons que l'ampleur de la baisse du crédit double lorsque les comptes d'une entreprise sont réaffectés d'une succursale spécialisée dans son secteur à une succursale qui ne l'est pas. L'ampleur de cet effet estimé est robuste au contrôle de la variation de la distance associée à la fermeture de la succursale. Nous constatons également que la baisse du crédit après la fermeture d'une succursale s'explique entièrement par la variation de la spécialisation sectorielle entre les succursales lorsque la nouvelle succursale est située dans une zone caractérisée par un niveau élevé de concurrence bancaire. Les résultats suggèrent fortement l'existence d'un marché du crédit bancaire segmenté, où la spécialisation des banques par industrie augmente le coût de substitution des sources de financement bancaire pour les petites entreprises.

Chapter 3: Implications macroéconomiques des flux de relations bancaires

Dans ce troisième chapitre, co-écrit avec Yasser Boualam, nous étudions le rôle des imperfections de marché dans les fluctuations du crédit aux entreprises, d'un point de vue

macroéconomique. Quels sont les moteurs des fluctuations du crédit au cours du cycle économique et à long terme ? Comment les banques ajustent-elles leur offre de crédit en réponse aux chocs globaux ou aux changements de politique ? Ces questions sont au premier plan de la recherche macro-financière et bancaire au moins depuis les travaux précurseurs de Bernanke (1983). Pourtant, notre compréhension des fluctuations globales du crédit et de leurs implications pour l'économie réelle reste largement incomplète.

Le crédit bancaire est une source de financement importante pour la majorité des entreprises. Un aspect particulièrement important qui a été largement étudié au niveau microéconomique, mais négligé au niveau macroéconomique, concerne les relations de crédit entre les banques et les entreprises. En effet, une vaste littérature théorique et empirique souligne le rôle de ces relations qui permettent d'atténuer les frictions d'agence et affectent ainsi l'allocation du crédit.[2] Cette littérature a également souligné l'existence d'une hétérogénéité en termes de qualité de l'appariement et des caractéristiques inhérentes à la relation, telles que la durée, qui peut potentiellement entraver la capacité des banques à ajuster leur offre de crédit sans friction (Boualam (2018)). À l'inverse, la plupart des modèles macrofinanciers supposent *simplement* que les emprunteurs et/ou les prêteurs sont homogènes, ou font abstraction de la nature de long-terme des contrats financiers et des frictions de marché qui peuvent empêcher les banques de former ou de rompre sans coût ces relations de crédit. Ces modèles minimisent donc l'importance des relations bancaires et impliquent que les banques puissent rapidement ajuster le nombre de leurs emprunteurs en réponse à des chocs. Ils laissent également peu de place à l'analyse du processus de réallocation du crédit entre les appariements banque-entreprise et de sa dynamique tout au long du cycle.

Dans ce chapitre, nous proposons une nouvelle perspective macroéconomique sur le processus d'intermédiation du crédit. Nous tentons de fournir des évidences empiriques nouvelles sur les rôles clés et distincts que jouent les marges *intensive* et *extensive* dans les fluctuations agrégées du crédit. Nous essayons ainsi de répondre à des questions simples, mais de premier ordre telles que : (i) Lorsque le crédit bancaire agrégé diminue de 5%, est-ce parce que la taille moyenne des prêts (c'est-à-dire la marge intensive) diminue de 5%, ou est-ce parce que 5% des appariements banque-entreprise (c'est-à-dire la marge extensive) sont détruits ? (ii) L'origine des fluctuations du crédit agrégé a-t-elle une importance ? (iii) Les chocs de politique monétaire ont-ils un impact différent sur ces marges ?

À notre connaissance, nous sommes les premiers à montrer que les banques ajustent activement à la fois le nombre *et* l'intensité de leurs relations, en réponse à des chocs macro économiques, et que ces deux marges représentent une source significative de la variation du

[2] Voir Boot (2000) et Degryse et al. (2009) pour un aperçu des travaux antérieurs.

total des prêts bancaires. Ces ajustements sont analogues à la manière dont les entreprises ajustent constamment à la fois la quantité d'heures travaillées et l'emploi. Ce point de vue peut sembler intuitif, pourtant — et de façon surprenante — une analyse approfondie de la dynamique de ces marges et de leurs implications macroéconomiques reste limitée, voire totalement absente de la littérature. En outre, nous établissons non seulement l'importance quantitative de ces marges, mais nous soutenons également qu'elles sont soumises à des comportements agrégés très différents. Ainsi, démêler les effets associés à chaque marge peut s'avérer instructif sur les mécanismes économiques en jeu et le rôle de la réallocation du crédit, et finalement produire des implications politiques pertinentes.

Pour faire la lumière sur ce processus, nous nous appuyons sur une source d'information essentielle, le registre français des crédits, qui couvre le marché des prêts commerciaux en France, et qui est maintenu par la Banque de France. Les données granulaires et quasi exhaustives permettent de suivre les appariements entre banques et entreprises et les expositions au crédit correspondantes sur la période 1998-2018. Pour étudier les propriétés des flux de relations de crédit, nous développons une méthodologie empirique semblable à celle mise en place par Davis and Haltiwanger (1992) pour les flux de travailleurs. Notre méthodologie prend en considération les caractéristiques spécifiques associées à la structure du marché du crédit et les limitations des données disponibles. Par exemple, nous suivons les entrées et les sorties de chaque correspondance banque-entreprise afin de déterminer le moment de la création et le moment supposé de la destruction de la relation, afin de construire des flux bruts de relations de crédit. Nous tenons également compte de l'hétérogénéité en coupe des caractéristiques des relations et des contrats financiers (la taille du prêt, le type et la maturité du crédit, et la durée de la relation).

Comprendre les implications macroéconomiques des flux de relations de crédit entre banques et entreprises est, comme nous l'avons montré, une entreprise assez naturelle. L'absence de faits empiriques sur le sujet découle directement d'un manque des données microéconomiques, en particulier sur une période suffisamment longue. Les études antérieures, telles que Dell'Ariccia and Garibaldi (2005), reposaient sur des données agrégées au niveau des banques. Elles ne permettent donc pas d'identifier les emprunteurs, ni observer les flux de relations bancaires (ou seulement les flux nets, au niveau de la banque). Par conséquent, ces études ne peuvent pas distinguer les marges extensives des marges intensives, ni saisir précisément l'ampleur et les propriétés sous-jacentes de la réallocation du crédit. Au contraire, nous proposons ici une nouvelle approche pour exploiter les informations disponibles dans les registres de crédit, qui sont généralement utilisées dans des contextes microéconomiques, afin de mettre en exergue de nouveaux résultats agrégés. Notre recherche établit les faits stylisés

suivants concernant les marges extensive et intensive du crédit :

1. Les marges extensives et intensives du crédit fluctuent continuellement dans le temps. Si leur tendance de long-terme est à peu près identique, la volatilité de la marge intensive est relativement plus élevée.
2. Les deux marges sont importantes pour expliquer les fluctuations du crédit au cours d'un cycle économique, la marge extensive contribuant entre un quart et la moitié de la variance du crédit agrégé.
3. A long terme, la marge extensive représente l'essentiel des variations du crédit global.

Notre analyse met également en évidence les caractéristiques suivantes concernant les flux bruts des relations de crédit :

1. La création, la destruction et la réaffectation des relations banques-entreprises coexistent tout au long du cycle.
2. La création (flux entrants) et la destruction (flux sortants) des relations présentent une plus grande volatilité par rapport aux flux nets. Les variations des flux nets sont principalement dues aux flux entrants.
3. Les sorties sont plus volatiles pour les prêts de faible montant, les prêts à court terme et les relations de crédit d'une durée inférieure à un an. Les entrées sont plus volatiles pour les relations avec de petits prêts ou des lignes de crédit.

Nos résultats soulignent également que les fluctuations (la baisse) du crédit aux entreprises observées pendant, ou à la suite, d'une crise économique sont causées conjointement par les marge extensives *et* intensives, ce qui suggère que différents mécanismes économiques peuvent être en jeu. Une meilleure compréhension de l'origine (marge extensive/intensive) d'une baisse du crédit peut donc être utile pour la conception d'outils de politique économique efficaces et ciblés. Pour approfondir ce point, nous étudions comment la politique monétaire se transmet par les canaux de la marge extensive et intensive. Nous montrons que si la marge intensive réagit immédiatement et fortement aux surprises de la politique monétaire, la réponse de la marge extensive est relativement plus graduelle et plus modérée. Nous notons également que le canal de la marge extensive opère principalement pour les banques relativement petites ou celles dont le portefeuille de prêts le plus flexible (i.e., des prêts à court-terme).

Finalement, notre cadre empirique fournit des outils pour mieux comprendre le processus de réallocation qui se produit sur les marchés du crédit et les canaux par lesquels les chocs bancaires se transmettent à l'économie réelle. En particulier, nous montrons que le taux de réallocation des relations de crédit est contracyclique, conformément à l'idée de *cleansing*

effect qui se produit pendant les récessions. En outre, nous montrons les taux annuels de réallocation (excédentaire) ont diminué de manière constante au cours des deux dernières décennies. Ces résultats indiquent l'existence de facteurs entravant la fluidité du marché du crédit.

Chapitre 4: Technologies de l'information et de la communication et commerce international : le cas du déploiement de l'ADSL en France

Dans ce dernier chapitre, co-écrit avec Clément Malgouyres et Thierry Mayer, nous étudions le rôle de l'Internet haut-débit dans la réduction des frictions de recherche auxquelles sont confrontés les importateurs français.

Pour cela, nous documentons l'effet de la diffusion des technologies de l'information et de la communication en France, entre 1997 et 2007, sur le commerce international de biens et sur le bien-être des consommateurs. Pour cela, nous utilisons le déploiement progressif de l'Internet haut-débit (ADSL) en France pour évaluer l'impact causal de l'accès à cette technologie sur les importations des entreprises françaises.

En pratique, nous mobilisons des données sur les variations spatiales et temporelles d'accès au haut débit entre communes pour comparer des entreprises bénéficiant d'un accès à celles n'en bénéficiant pas. Nous trouvons que l'expansion de l'ADSL entraîne une augmentation de la valeur des importations de 25% au niveau de l'entreprise et que cet effet positif passe principalement par la marge extensive (l'augmentation de la diversité des produits importés et du nombre de partenaires commerciaux).

Nous constatons également que l'augmentation des importations françaises est particulièrement pour des biens en provenance de Chine et d'Europe de l'Est et pour les biens d'équipement et les biens intermédiaires. Finalement, nous développons un modèle théorique, dans lequel les entreprises optimisent leur stratégie d'importation, qui nous permet de quantifier les effets de l'Internet haut-débit sur le bien-être des consommateurs. Dans le cadre de ce modèle, nos estimations impliquent que l'ADSL a réduit l'indice des prix à la consommation de 1,85% et que le canal des importations représente un quart de cet effet.

References

Aker, J. C. (2010). Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics*, 2(3):46–59.

Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Akerman, A., Leuven, E., and Mogstad, M. (2018). Information frictions, internet and the relationship between distance and trade. *mimeo*, page 48.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Bernanke, B. S. (1983). Non-monetary effects of the financial crisis in the propagation of the great depression. Technical report, NBER Working Paper.

Bhuller, M., Kostøl, A., and Vigtel, T. C. (2019). How Broadband Internet Affects Labor Market Matching. Memorandum 10/2019, Oslo University, Department of Economics.

Boot, A. W. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, 9(1):7–25.

Boualam, Y. M. (2018). Credit markets and relationship capital. *Working Paper at University of Pennsylvania,*.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634.

Davis, S. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, 107(3):819–863.

Degryse, H., Kim, M., and Ongena, S. (2009). *Microeconometrics of Banking Methods, Applications, and Results.* Oxford University Press.

Dell'Ariccia, G. and Garibaldi, P. (2005). Gross credit flows. *The Review of Economic Studies*, 72(3):665–685.

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory*, 3(2):156–168.

Duffie, D., Gârleanu, N., and Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, 73(6):1815–1847.

Eaton, J., Kortum, S., and Kramarz, F. (2018). Firm-to-Firm Trade: Imports, exports, and the labor market. Discussion papers 16048, Research Institute of Economy, Trade and Industry (RIETI).

Goyal, A. (2010). Information, direct access to farmers, and rural market performance in central india. *American Economic Journal: Applied Economics*, 2(3):22–45.

Hauswald, R. and Marquez, R. (2003). Information Technology and Financial Services Competition. *The Review of Financial Studies*, 16(3):921–948.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The Quarterly Journal of Economics*, 122(3):879–924.

Kiyotaki, N. and Wright, R. (1993). A search-theoretic approach to monetary economics. *The American Economic Review*, 83(1):63–77.

Kroszner, R. S. and Strahan, P. E. (1999). What drives deregulation? economics and politics of the relaxation of bank branching restrictions. *The Quarterly Journal of Economics*, 114(4):1437–1467.

Lendle, A., Olarreaga, M., Schropp, S., and Vézina, P.-L. (2016). There goes gravity: ebay and the death of distance. *Economic Journal*, 126(591):406–441.

Lenoir, C., Mejean, I., and Martin, J. (2018). Search Frictions in International Good Markets. Technical report.

Malgouyres, C. (2017). The impact of chinese import competition on the local structure of employment and wages: Evidence from france. *Journal of Regional Science*, 57(3):411–441.

McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126.

Mortensen, D. T. (1988). Matching: Finding a partner for life or otherwise. *American Journal of Sociology*, 94:S215–S240.

Paravisini, D., Rappoport, V., and Schnabl, P. (2017). Specialization in Bank Lending: Evidence from Exporting Firms. (12156).

Petersen, M. and Rajan, R. (2002). Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570.

Petersen, M. A. and Rajan, R. (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of finance*, 47(4):1367–1400.

Rauch, J. E. (2001). Business and social networks in international trade. *Journal of Economic Literature*, 39(4):1177–1203.

Rogerson, R., Shimer, R., and Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988.

Santos Silva, J. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Santos Silva, J. and Tenreyro, S. (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112(2):220–222.

Shimer, R. and Smith, L. (2000). Assortative matching and search. *Econometrica*, 68(2):343–370.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Steinwender, C. (2018). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–96.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Stiglitz, J. (1975). The theory of "screening," education, and the distribution of income. *American Economic Review*, 65(3):283–300.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410.

Udell, G. F. (2015). SME Access to Intermediated Credit: What Do We Know and What Don't We Know? In Moore, A. and Simon, J., editors, *Small Business Conditions and Finance*, RBA Annual Conference Volume (Discontinued). Reserve Bank of Australia.

Weill, P.-O. (2007). Leaning against the wind. *The Review of Economic Studies*, 74(4):1329–1354.

Zenou, Y. (2009). *Urban Labor Economics.* Cambridge University Press.